



(12) 发明专利申请

(10) 申请公布号 CN 104823188 A

(43) 申请公布日 2015. 08. 05

(21) 申请号 201380062686. 0

代理人 梁丽超 陈鹏

(22) 申请日 2013. 11. 26

(51) Int. Cl.

(30) 优先权数据

G06F 17/30(2006. 01)

13/689, 969 2012. 11. 30 US

(85) PCT国际申请进入国家阶段日

2015. 05. 29

(86) PCT国际申请的申请数据

PCT/US2013/072077 2013. 11. 26

(87) PCT国际申请的公布数据

W02014/085495 EN 2014. 06. 05

(71) 申请人 脸谱公司

地址 美国加利福尼亚州

(72) 发明人 尹戈尔·卡比廖 亚历山大·伊利克

花鸣 颜宏

(74) 专利代理机构 北京康信知识产权代理有限

责任公司 11240

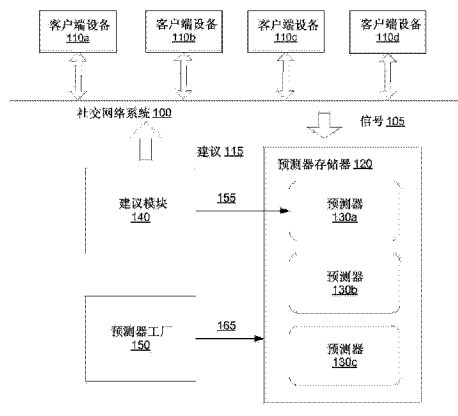
权利要求书4页 说明书16页 附图7页

(54) 发明名称

针对在线系统中的用户动作的定制化预测器

(57) 摘要

在线系统生成用于预测在线系统的用户动作的预测器。在线系统接收请求以生成用于预测用户是否可能采取特定动作类型的动作的预测器模型。请求规定动作类型和用于识别动作类型的成功实例和动作类型的失败实例的标准。在线系统收集包括动作类型的成功和失败实例的数据。在线系统使用生成的数据生成不同类型的一个或多个预测器。在线系统评估并且比较生成的不同预测器的性能并且基于性能选择预测器。在线系统将用于访问生成的预测器的句柄返回至预测器的请求者。



1. 一种具体旨在加快在线系统内的潜在用户交互的计算机实施的方法,包括:

通过在线系统接收用于生成预测器的多个请求,每个请求规定由所述在线系统的用户执行的动作类型和用于识别所述动作类型的成功实例的标准;

为与动作类型相关联的每个请求,通过所述在线系统生成针对所述动作类型的预测器,所述生成包括:

识别训练数据组,所述训练数据组包括由所述在线系统的用户执行的所述动作类型的多个成功实例和多个失败实例;

使用所识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型;

通过对包括由所述在线系统的用户执行的所述动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器,来评估每个候选预测器的性能;以及

通过比较所述一个或多个候选预测器的性能来选择用于所述动作类型的预测器;以及将与所生成的预测器对应的句柄返回至所选择的预测器的请求者。

2. 根据权利要求 1 所述的计算机实施的方法,进一步包括:

从在所述在线系统中运行的处理接收针对预测器的请求,所述请求规定用户信息的句柄和实例;

对所述用户信息的实例执行与所规定的句柄对应的所述预测器;以及

将所述预测器的结果返回至所述请求处理。

3. 根据权利要求 1 或 2 所述的计算机实施的方法,其中,动作类型的成功实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户采取该类型的动作的信息对应。

4. 根据权利要求 1 至 3 中任一项所述的计算机实施的方法,其中,所述请求进一步规定失败实例的标准,其中,动作类型的失败实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户未能采取该类型的动作的信息对应,

其中,优选地,动作类型的失败实例与存储在所述在线系统中的指示用户提供表示不喜欢在与该动作类型相关联的用户界面中呈现的信息的输入的信息对应。

5. 根据权利要求 1 至 4 中任一项所述的计算机实施的方法,其中,每种动作类型针对在所述在线系统中表示的实体,

其中,优选地,与所述动作类型对应的所述实体是所述在线系统的用户、社交群、或图像之一。

6. 根据权利要求 1 至 5 中任一项所述的计算机实施的方法,其中,候选预测器是基于线性回归、逻辑回归、提升树、或权重决策树之一的类型的机器学习模型。

7. 根据权利要求 1 至 6 中任一项所述的计算机实施的方法,进一步包括,针对动作类型的每个生成的预测器:

将所生成的预测器的性能与该动作类型的现有预测器的性能进行比较;以及

响应于所生成的预测器的性能超过所述现有预测器的性能,利用所生成的预测器替换所述现有预测器,

优选地,进一步包括,针对每种动作类型,定期地生成新预测器:

将所述新预测器的性能与之前生成的预测器的性能进行比较；以及
响应于所述新预测器的性能超过所述之前生成的预测器的性能，利用所述新预测器替换所述之前生成的预测器，

其中，优选地，基于所述新预测器的性能与之前生成的预测器性能的比较，确定定期生成针对每种动作类型的新预测器的速率，

其中，优选地，响应于所述新预测器的性能相比于所述现有预测器的性能的改善低于阈值，降低定期生成针对每种动作类型的新预测器的速率。

8. 一种具体根据权利要求 1 至 7 中任一项所述的计算机实施的方法，包括：

通过在线系统存储多个预测器，每个预测器用于预测所述在线系统的用户是否执行某一动作类型，每个预测器与用于确定所述动作类型的成功实例的标准相关联；

选择与动作类型对应的现有预测器；

识别数据组，所述数据组包括由所述在线系统的用户执行的所述动作类型的成功实例和失败实例；

使用所识别的数据组训练一个或多个候选预测器，其中，每个候选预测器包括机器学习模型；

通过对包括由所述在线系统的用户执行的所述动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器，来评估每个候选预测器的性能；

通过比较所述一个或多个候选预测器的性能从所述一个或多个候选预测器中选择候选预测器；

将所选择的候选预测器的性能与所述现有预测器的性能进行比较；以及

响应于所选择的候选预测器的性能超过所述现有预测器的性能，由所选择的候选预测器替换所述现有预测器。

9. 根据权利要求 8 所述的计算机实施的方法，其中，动作类型的成功实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户采取该类型的动作的信息对应，

其中，优选地，动作类型的失败实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户未能采取该类型的动作的信息对应，

其中，优选地，候选预测器是基于线性回归、逻辑回归、提升树、或权重决策树之一的类型的机器学习模型。

10. 根据权利要求 9 所述的计算机实施的方法，进一步包括，定期地生成新预测器：

将所述新预测器的性能与之前生成的预测器的性能进行比较；以及

响应于所述新预测器的性能超过所述之前生成的预测器的性能，利用所述新预测器替换所述之前生成的预测器，

其中，优选地，响应于所述新预测器的性能相比于所述现有预测器的性能的改善低于阈值，降低定期生成针对每种动作类型的新预测器的速率。

11. 一个或多个计算机可读非瞬时性存储介质，包含软件，所述软件可操作为当运行时执行根据权利要求 1 至 10 中任一项所述的方法。

12. 一种系统包括：一个或多个处理器；以及存储器，耦接至所述处理器，所述存储器

包括由所述处理器可运行的指令,所述处理器可操作为当运行所述指令时执行根据权利要求 1 至 10 中任一项所述的方法。

13. 一种计算机实施的方法,包括:

通过在线系统存储多个预测器,每个预测器用于预测所述在线系统的用户是否执行某一动作类型,每个预测器与用于确定所述动作类型的成功实例的标准相关联;

选择与动作类型对应的现有预测器;

识别数据组,所述数据组包括由所述在线系统的用户执行的所述动作类型的成功实例和失败实例;

使用所识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型;

通过对包括由所述在线系统的用户执行的所述动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器,来评估每个候选预测器的性能;

通过比较所述一个或多个候选预测器的性能从所述一个或多个候选预测器中选择候选预测器;

将所选择的候选预测器的性能与所述现有预测器的性能进行比较;以及

响应于所选择的候选预测器的性能超过所述现有预测器的性能,由所选择的候选预测器替换所述现有预测器。

14. 根据权利要求 13 所述的计算机实施的方法,其中,动作类型的成功实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户采取该类型的动作的信息对应。

15. 根据权利要求 13 所述的计算机实施的方法,其中,动作类型的失败实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户未能采取该类型的动作的信息对应。

16. 根据权利要求 13 所述的计算机实施的方法,其中,候选预测器是基于线性回归、逻辑回归、提升树、或权重决策树之一的类型的机器学习模型。

17. 根据权利要求 13 所述的计算机实施的方法,进一步包括,定期地生成新预测器:

将所述新预测器的性能与之前生成的预测器的性能进行比较;以及

响应于所述新预测器的性能超过所述之前生成的预测器的性能,利用所述新预测器替换所述之前生成的预测器。

18. 根据权利要求 17 所述的计算机实施的方法,其中,响应于所述新预测器的性能相比于所述现有预测器的性能的改善低于阈值,降低定期生成针对每种动作类型的新预测器的速率。

19. 一种具有非瞬时性计算机可读存储介质的计算机程序产品,所述非瞬时性计算机可读存储介质存储计算机可执行代码,所述代码包括:

在线系统的预测器工厂模块,被配置为:

存储多个预测器,每个预测器用于预测所述在线系统的用户是否执行某一动作类型,每个预测器与用于确定所述动作类型的成功实例的标准相关联;

选择与动作类型对应的现有预测器;

识别数据组,所述数据组包括由所述在线系统的用户执行的所述动作类型的成功实例

和失败实例；

使用所识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型；

通过对包括由所述在线系统的用户执行的所述动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器,来评估每个候选预测器的性能；

通过比较所述一个或多个候选预测器的性能从所述一个或多个候选预测器中选择候选预测器；

将所选择的候选预测器的性能与所述现有预测器的性能进行比较；以及

响应于所选择的候选预测器的性能超过所述现有预测器的性能,由所选择的候选预测器替换所述现有预测器。

20. 根据权利要求 19 所述的计算机程序产品,其中,动作类型的成功实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户采取该类型的动作的信息对应。

21. 根据权利要求 19 所述的计算机程序产品,其中,动作类型的失败实例与存储在所述在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面所述用户未能采取该类型的动作的信息对应。

针对在线系统中的用户动作的定制化预测器

技术领域

[0001] 本发明涉及预测例如社交网络系统的在线系统中的用户动作,并且具体地涉及针对在线系统中的用户动作的定制化预测器的生成,特别地旨在加快在线系统内的潜在用户交互。

背景技术

[0002] 在线系统通常向用户呈现有用的信息并且允许用户与在线系统交互。在线系统可使用技术以在向用户呈现信息之前确定用户可能感兴趣的信息。如果给他们呈现他们喜欢的信息,则用户更可能定期地访问在线系统。在线系统经常从广告赚取收益。广告商更喜欢在由他们的用户定期访问的在线系统中做广告。因此,用户忠实度可确定使用在线系统生成的收益。因此,在线系统向用户呈现感兴趣信息的能力通常影响通过在线系统赚取的收益。

[0003] 一些系统使用机器学习技术确定与单个用户相关的信息以预测用户动作。例如,如果机器学习模型预测用户可能观看某些内容,则在线系统可选择向用户呈现该内容而不是其他内容。另一方面,如果机器学习模型预测用户可能对某些信息不感兴趣,则在线系统可选择向用户呈现的一些其他信息。

[0004] 然而,开发机器学习模型需要在线系统的开发人员的大量努力。通常,机器学习模型的开发人员指定使用的机器学习技术、与机器学习模型相关的各种特征,并且提供用于训练机器学习模型的训练数据。在复杂的在线系统中,开发人员可能不知道各种特征对预测器模型的影响。通常,开发人员甚至可能不知道能够在模型中使用的复杂系统中可获得的所有可能特征。此外,各种特征对模型的影响可随着时间而改变。在技术方面的新开发可能以不可预测的方式影响各种特征对模型的影响。例如,虽然未结合在线系统开发新技术,但是如果新类型的移动技术是可获得的,则系统的用户可以表现得不同。在用户行为方面的变化能够改变各种特征对模型的影响。当考虑到他们自己在线系统的所有可能方面以及能够影响模型的外部因素时,这对于机器学习模型的开发人员开发并且维持模型是困难的。因此,传统技术给机器学习模型的开发人员施加沉重负担。

发明内容

[0005] 本发明的实施方式允许在线系统生成用于预测在线系统的用户动作的模型。例如,诸如社交网络系统的在线系统可使用模型预测用户是否可能建立与另一用户的联系、用户是否可能评论图像、或用户是否可能加入社交群。这些模型在本文中也被称为预测器。社交网络系统使用预测器选择向社交网络系统的用户呈现的信息,例如,提出与用户的潜在联系、选择向用户呈现的内容、或向用户推荐社交群。

[0006] 在实施方式中,在线系统接收请求以生成用于预测用户是否可能采取特定动作类型的动作的新预测器。例如,动作类型可对应于用户喜欢图像、用户发送请求与另一用户联系、或用户加入社交群。请求指定动作类型和用于识别动作类型的成功或失败的标准。例

如,如果动作类型对应于用户发送请求以与另一用户联系,则动作的成功实例(successful instance,成功情形)对应于用户发送请求以与由在线系统提出的潜在朋友联系,并且失败实例对应于用户不能联系到潜在朋友或试图隐藏朋友建议。在线系统从用户在一段时间中与在线系统的交互中收集包括动作类型的成功和失败实例的数据。在线系统使用生成的数据生成不同类型的一个或多个预测器。例如,一个预测器可使用线性回归(linear regression),另一预测器可以使用逻辑回归(logistic regression),并且第三预测器可以使用提升树(boosting tree)或权重决策树。在线系统评估并且比较生成的不同预测器的性能并且基于性能选择预测器。在线系统将用于访问生成的预测器的句柄(handle)返回至预测器的请求者。因此,请求方接收针对动作的定制化预测器而不用了解关于机器学习的任何事情。

[0007] 在实施方式中,在线系统基于通过在线系统收集的新数据再生成新预测器以替换现有预测器。新预测器能够可以是使用完全不同的机器学习技术的新模型、或使用相同机器学习技术但是具有在模型内使用的新系数值的模型。可以预定速率定期地生成新预测器。可自适应地修改生成新预测器的速率。例如,如果在新预测器中观察的性能相比于针对动作类型的现有预测器的改善低于阈值,则可降低针对该具体动作类型的预测器的再生成速率。

[0008] 在本发明内容部分和以下详细说明中所述的特征和优势不是全部包含的。鉴于附图、说明书、和权利要求,很多其他特征和优点对于本领域普通技术人员将是显而易见的。

[0009] 涉及方法、存储介质和系统的所附权利要求中具体地公开了根据本发明的实施方式,其中,在一个权利要求类型(例如,方法)中提及的任何特征也能够另一权利要求类型(例如,系统)中要求保护。

[0010] 在实施方式中,根据本发明,提供了特别旨在加快在线系统内的潜在用户交互的计算机实施的方法,包括:

[0011] 通过在线系统接收用于生成预测器的多个请求,每个请求指定通过在线系统的用户执行的动作类型和用于识别动作类型的成功实例的标准;

[0012] 为与动作类型相关联的每个请求,通过在线系统生成用于该动作类型的预测器,该生成包括:

[0013] 识别包括通过在线系统用户执行的动作类型的多个成功实例和多个失败实例的训练数据组;

[0014] 使用识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型;

[0015] 通过对包括由在线系统的用户执行的动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器来评估每个候选预测器的性能;以及

[0016] 通过比较一个或多个候选预测器的性能选择用于动作类型的预测器;以及

[0017] 将与生成的预测器对应的句柄返回至所选择的预测器的请求者。

[0018] 计算机实施的方法可进一步包括:

[0019] 接收来自在线系统中运行的处理的针对预测器的请求,请求指定用户信息的实例和句柄;

[0020] 对用户信息的实例执行与指定的句柄对应的预测器;以及

- [0021] 将预测器的结果返回至请求处理。
- [0022] 优选地,动作类型的成功实例对应于存储在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的用户界面用户采取该类型的动作的信息。
- [0023] 具体地,请求进一步指定针对失败实例的标准,其中,动作类型的失败实例对应于存储在线系统中的指示响应于给用户呈现允许用户采取该动作类型的动作的信息用户未采取该类型的动作的信息。
- [0024] 动作类型的失败实例可对应于存储在线系统中的指示用户提供表示不喜欢在与动作类型关联的用户界面中呈现的信息的输入的信息。
- [0025] 优选地,每种类型的动作瞄准在线系统中表示的实体。
- [0026] 具体地,对应于动作类型的实体是在线系统的用户、社交群、或图像之一。
- [0027] 候选预测器能够是基于线性回归 (linear regression)、逻辑回归 (logistic regression)、提升树 (boosting tree)、权重决策树之一的机器学习模型的类型。
- [0028] 计算机实施的方法进一步可包括为动作类型的每个生成的预测器:
- [0029] 将生成的预测器的性能与该类型动作的现有预测器的性能比较;以及
- [0030] 响应于生成的预测器的性能超过现有预测器的性能,利用生成的预测器替换现有预测器。
- [0031] 计算机实施的方法可进一步包括,为每个动作类型,定期地生成新预测器:
- [0032] 将新预测器的性能与之前生成的预测器的性能进行比较;以及
- [0033] 响应于新预测器的性能超过之前生成的预测器的性能,利用新预测器替换之前生成的预测器。
- [0034] 优选地,基于新预测器的性能与之前生成的预测器的性能的比较确定定期生成用于每种动作类型的新预测器的速率。
- [0035] 具体地,响应于新预测器的性能相比现有预测器的性能的改善低于阈值降低定期为每种动作类型生成新预测器的速率。
- [0036] 在也能够要求保护的又一实施方式中,提供了计算机实施的方法,包括:
- [0037] 通过在线系统存储多个预测器,每个预测器用于预测在线系统的用户是否执行某种类型的动作,每个预测器与用于确定该类型动作的成功实例的标准相关联;
- [0038] 选择与某一类型的动作对应的现有预测器;
- [0039] 识别包括通过在线系统用户执行的动作类型的成功实例和失败实例的数据组;
- [0040] 使用识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型;
- [0041] 通过对包括通过在线系统用户执行的动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器来评估每个候选预测器的性能;
- [0042] 通过比较一个或多个候选预测器的性能从一个或多个候选预测器中选择候选预测器;
- [0043] 将所选择的候选预测器性能与现有预测器性能比较;以及
- [0044] 响应于所选择的候选预测器的性能超过现有预测器的性能,通过所选择的候选预测器替换现有预测器。
- [0045] 优选地,动作类型的成功实例对应于存储在在线系统中的指示响应于给用户呈现

允许用户采取该动作类型的动作的用户界面用户采取该类型的动作的信息。

[0046] 优选地,动作类型的失败实例对应于存储在在线系统中的指示响应于给用户呈现允许用户未能采取该动作类型的动作的用户界面用户未能采取该类型的动作的信息。

[0047] 计算机实施的方法可以进一步包括,定期地生成新预测器:

[0048] 将新预测器的性能与之前生成的预测器的性能比较;以及

[0049] 响应于新预测器的性能超过之前生成的预测器的性能,利用新预测器替换之前生成的预测器。

[0050] 优选地,其中,响应于新预测器的性能相比于现有预测器的性能的改善低于阈值,降低定期生成针对每种动作类型的新预测器的速率。

[0051] 在根据本发明也能够要求保护的另—实施方式中,提供了具有存储计算机可执行代码的非瞬时性计算机可读存储介质的计算机程序产品,代码包括:

[0052] 在线系统的预测器工厂模块被配置为:

[0053] 存储多个预测器,每个预测器用于预测在线系统的用户是否执行某一类型的动作,每个预测器与用于确定该类型的动作的成功实例的标准相关联;

[0054] 选择与某一类型的动作对应的现有预测器;

[0055] 识别包括通过在线系统用户执行的动作类型的成功实例和失败实例的数据组;

[0056] 使用识别的数据组训练一个或多个候选预测器,其中,每个候选预测器包括机器学习模型;

[0057] 通过对包括通过在线系统的用户执行的动作类型的多个成功实例和多个失败实例的测试数据组执行每个预测器来评估每个候选预测器的性能;

[0058] 通过比较一个或多个候选预测器的性能从一个或多个候选预测器中选择候选预测器;

[0059] 将所选择的候选预测器的性能与现有预测器的性能比较;以及

[0060] 响应于所选择的候选预测器的性能超过现有预测器的性能,通过所选择的候选预测器替换现有预测器。

[0061] 在也能够要求保护的发明的又一实施方式中,一个或多个计算机可读非瞬时性存储介质包含当执行时可操作为执行根据本发明的方法或任何以上提及的实施方式的软件。

[0062] 在也能够要求保护的发明的又一实施方式中,一种系统包括:一个或多个处理器;以及存储器,耦接至处理器,存储器包括通过处理器可执行的指令,当执行指令时处理器可操作为执行根据本发明的方法或任何以上提及的实施方式。

附图说明

[0063] 图 1A 是根据本发明的实施方式的用于在例如社交网络系统的在线系统中生成并且调用定制化预测的系统环境的示图。

[0064] 图 1B 是示出根据本发明的实施方式的用于由用户采取的特定动作的预测器的实例的示图。

[0065] 图 2A 示出了根据本发明的实施方式的允许为由用户采取的特定动作生成和调用预测器的例如社交网络系统的在线系统的系统架构。

[0066] 图 2B 是根据本发明的实施方式的在线系统的预测器工厂模块 (factory module) 的系统架构的示图。

[0067] 图 3 示出了根据本发明的实施方式用于生成由社交网络系统中的各种模块使用的预测器模型的整体处理。

[0068] 图 4 示出了根据本发明的实施方式为由用户的给定类型动作生成预测器模型的整体处理。

[0069] 图 5 是根据本发明的一个实施方式用于在社交网络系统中再生成模型并且更新它们的处理的流程图。

[0070] 附图仅出于说明性的目的描述了本发明的各种实施方式。本领域技术人员将从下列讨论中很容易认识到在不背离本公开中所描述的的原理的情况下,可采用本公开中示出的结构和方法的替换实施方式。

具体实施方式

[0071] 现在将详细地参考若干实施方式,在附图中示出其实例。应注意,在可行的情况下类似或相同的参考标号可在附图中使用并且可表示类似或相同的功能。附图仅出于说明性的目的描述了所公开系统(或方法)的实施方式。本领域技术人员从下列描述中容易认识到,在不背离本文中所述的原理的情况下,可采用本文中所示出的结构和方法的替换实施方式。

[0072] 系统环境

[0073] 图 1A 是根据本发明的实施方式的用于生成并且调用例如社交网络系统的在线系统中的定制化预测器的系统环境的示图。虽然使用社交网络系统示出,但是本文中论述的本发明可应用于允许用户与在线系统交互的任何在线系统。具体地,社交网络系统为其用户提供与社交网络系统的其他用户通信和交互的能力。用户加入社交网络系统并且然后添加与他们希望联系的许多其他用户的联系。如本文中使用的,术语“朋友”指代用户经由社交网络系统与其形成联系、关联或关系的任何其他用户。

[0074] 图 1 和其他附图使用相同参考标号以识别相同元件。在参考标号之后的字母(诸如“110a”)表示该文本具体指具有该给定参考标号的元件。诸如“110”的文本中没有后跟字母的参考标号指附图中具有那个参考标号的任何或所有元件(例如,文本中“110”指附图中的参考标号“110a”和/或“110b”)。

[0075] 用户使用客户端设备 110 与社交网络系统 100 交互。在一个实施方式中,客户端设备 110 可以是运行例如微软的 Windows- 兼容式操作系统(OS)、苹果的 OS X、和/或 Linux 发行版本的操作系统个人计算机(PC)、台式计算机、便携式计算机、笔记本、平板电脑。在另一实施方式中,客户端设备 110 可以是具有计算机功能的任何设备,诸如个人数字助理(PDA)、移动电话、智能手机等。

[0076] 社交网络系统 100 向用户呈现不同类型的信息,例如,描述其他用户、社交群、社交活动、内容、图像等的信息。通常存在与用户相关的发生在社交网络系统中的大量动作。因为用户通常具有有限的时间花费在社交网络系统上并且在社交网络系统的用户界面中可获得的空间量也受限制,所以社交网络系统选择用户最感兴趣的信息。社交网络系统可结合从各种可获得的选项中选择用于呈现给用户的信息的一个或多个建议模块 140。

[0077] 建议模块 140 使用预测器 130 确定用户是否可能基于呈现给用户的信息执行希望的动作。例如,预测器 130 可确定用户请求与向用户呈现的新鲜事相关的更多信息的可能性。或预测器 130 可确定用户评论向用户呈现的图像的可能性。替换地,预测器 130 可确定发送与建议为用户的潜在联系的用户联系的请求的可能性。此外,可能存在不同类型的建议模块,每个使用一组不同的预测器。

[0078] 本发明的实施方式允许请求者在功能层面向预测器工厂 150 规定预测器的要求而不是要求请求者规定在模型中使用的各个特征。例如,请求者可规定预测器模型针对建议模块 140 需要的动作类型。例如,动作类型可以是当呈现一系列潜在联系时用户发送请求建立与另一用户的联系,或者动作类型可以是当呈现图像时用户评论图像。请求者还规定识别规定类型的动作何时能够视为成功或失败的标准。在实施方式中,每个请求器可提供能够使用应用程序编程接口 (API) 执行的一组可实行指令,当特定类型的动作在运行系统中发生时调用 API。可以基于对 API 的调用确定动作的成功。

[0079] 预测器工厂 150 使用通过请求者提供的信息生成预测器 130。因为可能是针对复杂系统的多个请求者,所以预测器工厂 150 可生成 165 若干个预测器 130a、130b、130c 等。预测器工厂 150 使用通过社交网络系统 100 的用户提供的信号 105 训练预测器模型。信号可对应于当用户与社交网络系统 100 交互时采取的各种动作。

[0080] 通过社交网络系统的模块调用生成的预测器 130。例如,建议模块 140 可以调用 155 预测器 130a,并且另一个模块可以调用 155 预测器 130b 等。例如建议模块 140 的社交网络系统的模块调用 155 合适的预测器以确定能够通过模块采取的各种动作,例如,以确定某些信息是否应当向具体用户呈现或建议模块 140 是否应当向用户提出某些建议 115。

[0081] 社交网络系统 100 包括在一个或多个计算机处理器上执行的软件模块。系统 100 和系统 110 的一些实施方式具有与本文所述的实施方式不同和 / 或其他的模块,并且可以不同于本文中描述的方式将功能分配在模块中。社交网络系统 100 可以包括除在图 1 中示出之外的模块,例如,本文中将进一步描述的在图 2 中示出的模块。

[0082] 图 1B 是示出根据本发明的实施方式的针对由用户采取的特定类型动作的预测器的实例的示意图。通过用户执行的动作可相对于目标实体。目标实体的实例包括在社交网络系统 100 内表示的其他用户、社交群、图像等。

[0083] 动作类型 185 可以取决于目标实体。例如,针对代表社交群的目标实体 180,动作 185 可以是源用户 170 加入社交群。针对代表社交网络系统用户的目标实体 180,动作 185 可以是源用户 170 发送消息、发送请求建立联系、或观看描述目标用户 180 的信息。针对代表图像的目标实体 180,动作 185 可以对应于源用户评论目标图像、喜欢目标图像、或观看目标图像。

[0084] 预测器工厂 150 可生成针对与具体类型目标实体相关的每种类型动作的预测器模型 130。例如,预测器 130p 预测类型 180a 的目标实体的动作类型 185a,预测器 130q 预测类型 180a 的目标实体的动作类型 185b,并且预测器 130r 预测类型 180b 的目标实体的动作类型 185c。在一些实施方式中,源可以是社交网络系统中表示的允许对另一目标实体执行动作的任何实体。允许实体执行的动作可不同于允许用户执行的动作。

[0085] 系统架构

[0086] 图 2A 是作为社交网络系统 100 的在线系统 100 的实施方式的系统架构的示意图。虽

然社交网络系统在本文中被描述为实例在线系统,但是本文中描述的原理可应用于其他在线系统。社交网络系统 100 包括预测器工厂 150、新鲜事生成器 235、网络服务器 215、动作记录器 240、动作日志 245、联系人存储器 230、用户个人资料存储器 225、以及建议模块 140。在其他实施方式中,社交网络系统 100 可包括用于各种应用的另外的、较少的或不同的模块。为了不使该系统的细节模糊,未示出诸如网络接口、安全机制、载荷平衡器、故障转移服务器、管理和网络操作控制台等的常规部件。

[0087] 社交网络系统 100 在用户个人资料存储器 225 中存储用户个人资料。用户个人资料存储器 225 存储描述社交网络系统 100 用户的信息,包括传记、人口统计以及其他类型的描述性信息,诸如工作经历、教育背景、性别、性倾向、爱好或偏好、位置等。用户个人资料存储器 225 也可存储通过用户提供的内容,例如,图像、视频、评论、和状态更新。在实施方式中,社交网络系统 100 的用户可以是一个组织,例如,商业、非盈利组织、制造商、供应商等。存储在组织的用户个人资料中的信息类型可不同于存储在个体用户个人资料中的信息。例如,组织可存储描述商业类型、与组织相关的财务信息、组织结构等。

[0088] 社交网络系统 100 允许用户添加与他们希望联系的社交网络系统 100 的多个其他用户的联系。可以由用户明确添加联系,例如,用户选择给定的其他用户成为其朋友,或基于用户的共同特征(例如,作为同一教育机构的男校友的用户)通过社交网络系统自动创建联系。社交网络系统可一起存储描述用户联系的信息以及特定于用户的信息。

[0089] 社交网络系统 100 在联系人存储器 230 中存储描述不同成员之间的一个或多个联系的数据。联系信息可以表示具有相似或共同工作经验、团体成员资格、爱好或教育背景的成员。此外,社交网络系统 100 包括用户定义的不同用户之间的联系,以允许用户规定与其他用户的关系。例如,这些用户定义联系允许成员与平行于该用户现实生活关系的诸如朋友、同事、伙伴等的其他用户生成关系。

[0090] 用户可从预定类型的联系中选择、或根据需要定义他们自身的联系类型。描述每个用户的用户信息可包括描述用户的联系的信息。此外,可相对于由用户执行的动作访问描述用户联系的信息。例如,如果用户在社交网络系统上张贴评论,则社交网络系统可将描述该动作的信息提供给用户的联系人。信息可经由新鲜事提供至用户的联系人。

[0091] 动作记录器 240 能够接收来自网络服务器 214 的关于登录和/或退出社交网络系统 100 的用户动作的通信。动作记录器 240 利用关于用户动作的信息填充动作日志 245 以追踪他们。通过在数据库或其他数据存储库(诸如动作日志 245)中维护的信息,特定用户相对于另一用户采取的任何动作与每个用户的个人资料相关。例如,这些动作可包括添加与其他用户的联系、向其他用户发送消息、读取来自其他用户的消息、观看与其他用户相关的内容、参加由另一用户发布的事件等。

[0092] 社交网络系统 100 维持向用户提供在社交网络系统 100 中可获得信息的定期更新的新鲜事渠道。通过新鲜事生成器 235 确定经由新鲜事渠道报告的信息。基于存储在动作日志 245 中的动作,新鲜事生成器 235 为每个用户生成关于可能与用户相关的信息的信息。这些消息被称为“故事(story)”;每个故事是包括基于与具体成员相关的动作日志中的一个或多个动作的一行或几行信息的信息。例如,如果用户的联系人执行交易,则动作可以经由新鲜事内容报告给用户。经由新鲜事报告的动作通常是通过用户的联系人执行的动作,但不限于这些。例如,如果与用户的联系人不相关的某些信息被确定为对用户有用,则信息

可经由新鲜事报告给用户。

[0093] 网络服务器 215 经由网络 210 将社交网络系统 100 链接至一个或多个客户端设备 110；网络服务器 215 用作网页、以及其他网站相关内容，诸如 Flash、XML 等。网络服务器 215 提供在社交网络系统 100 和客户端设备 110 之间接收并且路由消息的功能。这些消息能够是即时消息、队列消息（例如，电子邮件）、文本和 SMS（短消息服务）消息、或任何其他合适的消息技术。在一些实施方式中，社交网络系统 100 的其他用户能够看到由一个用户发送至另一用户的消息，例如，被接收消息的用户的联系人看到。除消息的接收者之外，社交网络系统 100 的其他用户能够看到的消息类型的实例是墙张贴（wall post）。墙张贴允许用户经由称为墙的通信渠道张贴消息，墙能够通过如通过用户隐私设置定义的一组用户访问。例如，用户能够与用户的所有联系人、与用户联系人的子组、与除了几个具体列出的联系人之外的所有联系人、或通过明确提供的联系人列表分享用户的墙。在一些实施方式中，用户可向另一用户发送仅可由另一用户访问的私人消息。

[0094] 社交网络系统 100 可为用户提供对由网站支持的各种类型的实体采取动作的能力。这些实体可包括网站成员可属于其的群或网络（其中，“网络”在此不仅指物理通信网络，而且指人们的社交网络）、成员可能感兴趣的事件或日历条目、成员可经由网站使用的计算机类应用、以及允许成员经由网站买、卖、拍卖、出租、或交换项目的交易。用户个人资料可存储用户与各种实体的关联。

[0095] 社交网络系统 100 可为用户提供用于彼此通信或获得他们感兴趣的信息（例如，他们的朋友参加的活动、他们的朋友安装的应用、由朋友对其他朋友的活动进行的评论等）的各种机制。在成员之间的通信机制被称为渠道。例如，如果用户与另一用户通信，则可能必需访问两个用户的用户信息以与发送者和接收者通信的动作相关联。

[0096] 建议模块 140 识别对各个用户感兴趣的信息并且将信息发送给他们。例如，社交网络系统可以向用户发送描述由与用户联系的其他用户采取的动作的故事。故事可经由例如新鲜事渠道的社交网络系统的通信渠道传送给用户。建议模块 140 使用在各个用户的用户个人资料中可获得的信息来确定每个用户感兴趣的内容。建议模块使用由预测器工厂 150 提供的预测器从可用于呈现的不同信息中选择呈现给用户的信息。预测器工厂 150 在图 1A 的描述中被描述，并且本文中更详细地被描述，例如在图 2B 中。

[0097] 客户端设备 110 运行浏览器 220 以允许用户与社交网络系统 100 交互。浏览器 220 允许用户使用社交网络系统 100 执行各种动作。这些动作包括检索用户感兴趣的信息、向其他用户推荐内容、将内容上传至社交网络系统 100、与社交网络系统的其他用户交互、建立与社交网络系统用户的联系等。

[0098] 通常经由网络 210（例如，经由因特网）执行在客户端设备 110 和在线系统 100 之间的交互。网络 210 实现在客户端设备 110 和在线系统 100 之间通信。在一个实施方式中，网络 210 使用标准通信技术和 / 或协议。因此，网络 210 可包括使用诸如以太网、802.11、微波存取全球互通（WiMAX）、3G、数字用户线路（DSL）、异步传输模式（ATM）、无限宽带、PCI 高速先进切换等的技术的链接。类似地，在网络 210 中使用的网络协议能够包括多协议标签交换（MPLS）、传输控制协议 / 因特网协议（TCP/IP）、用户数据报协议（UDP）、超文本传输协议（HTTP）、简单邮件传输协议（SMTP）、文件传输协议（FTP）等。可使用包括超文本标记语言（HTML）、可扩展标记语言（XML）等的技术和 / 或格式表示在网络 210 上交换的数据。此

外,可使用诸如安全套接字层 (SSL)、传输层安全性 (TLS)、虚拟专用网络 (VPNs)、因特网协议安全 (IPsec) 等的传统加密技术对所有或某些链接进行加密。在另一实施方式中,代替或除了以上描述的之外,实体可使用自定义的和 / 或专用的数据通信技术。根据实施方式,网络 210 也能够包括至其他网络 (诸如因特网) 的链接。

[0099] 图 2B 是根据本发明的实施方式的社交网络系统的预测器工厂模块 150 的系统架构的示意图。预测器工厂 150 包括包含机器学习模型 275、调度器 285、数据提取模块 265、模型评估器 280、模型生成器 290、动作数据存储器 260、以及预测器存储器 120 的模块。预测器工厂模块 150 接收请求以生成在社交网络系统内使用的预测器。在一些实施方式中,可提供由外部系统使用的预测器,例如,经由应用程序编程接口 (API) 调用在社交网络系统内的功能的外部系统。

[0100] 由预测器工厂模块 150 接收的对预测器的请求规定待预测的动作类型。动作类型可与社交网络系统内表示的具体类型实体相关。例如,可要求预测器预测源用户建立与目标用户联系的可能性。在这种情况下,动作是建立联系并且目标实体是另一个用户。如另一实例,可要求预测器预测源用户加入社交群的可能性。在这个实例中,动作是用户加入社交群并且目标实体是社交群。

[0101] 请求还规定用于识别动作的成功实例的标准和用于识别动作类型的失败实例的标准。例如,如果动作类型对应于用户发送与另一用户联系请求,则动作的成功实例对应于用户发送与由在线系统提出的潜在朋友联系请求,并且失败实例对应于用户未能与潜在朋友联系或试图隐藏朋友建议。用于识别成功实例的标准可包括指示向用户呈现的合适用户界面的信息和指示通过使用用户界面的用户执行的动作类型的信息。例如,可通过检测建议与用户的潜在联系人的用户界面呈现给用户、以及用户经由用户界面提供的输入指示用户想与建议的联系人之一联系的事实,来识别用户试图与另一用户建立联系的成功。

[0102] 某些动作类型的失败实例也可对应于用户提供负面反馈,例如,通过请求社交网络系统隐藏向用户呈现的某些信息。例如,社交网络系统 100 可将某些内容与允许用户提供关于该内容的反馈的用户界面一起推荐给用户。用户可使用用户界面以提供可以是正面或负面的反馈。负面反馈可以对应于用户经由用户界面指示用户不喜欢所呈现的信息、用户试图隐藏所呈现的信息,或指示用户不想再看见该具体类型的信息或来自该具体源的信息。

[0103] 可通过检测建议用户的潜在联系人的用户界面呈现给用户、以及未从用户接收到试图建立与建议的用户的联系的输入,来识别相同类型动作的失败。可通过在超过阈值时间间隔内未接收到试图建立联系的任何输入、或当替换先前的用户界面呈现新用户界面时或向用户呈现新的潜在联系人列表,来检测未从用户接收到输入。如果用户提供明确或含蓄的负面信号,则也可识别失败,例如,用户指示用户想要隐藏描述所建议的用户的信息,或用户指示用户不希望看见用户界面。可在运行时例如通过动作记录器、或通过处理在离线的社交网络系统中的动作 (例如通过分析动作日志) 来检测动作。

[0104] 数据提取模块 265 分析在社交网络系统中可获得的信息以提取用于训练、评估、或测试预测器模型的数据。例如,如果模块请求预测器工厂提供针对具体动作类型的预测器模型,则数据提取模块 265 提取该动作类型的实例。数据提取器 265 可通过分析在动作日志 245 中可获得的数据收集给定动作类型的数据。例如,数据提取器 265 可以寻找在动

作日志中给定动作类型的各种实例。

[0105] 例如,如果动作对应于用户评论照片,则动作的成功实例对应于当呈现照片时特定用户评论照片,并且失败实例对应于尽管呈现照片但是用户不评论照片。类似地,如果动作对应于用户请求针对新鲜事的额外信息,则成功实例对应于当呈现新鲜事时特定用户请求额外信息,并且失败实例对应于尽管呈现新鲜事但是用户不请求额外信息。数据提取器 265 识别动作的成功和失败实例两者,其中,正面成功对应于用户执行动作并且失败实例是在执行动作的合适机会下用户不执行动作。

[0106] 如果请求模块规定针对特定类型的用户请求预测器模型,则数据提取器 265 基于由属于该类型的用户采取的动作生成数据。请求者模块可请求针对具有特定人口统计特性的用户的预测器模型,例如,具有男性性别并且属于特定年龄群的用户、具有女性性别并且具有特定经济状况的用户、属于特定地理区域的用户、具有特定民族背景的用户、或讲特定语言的用户。

[0107] 在实施方式中,数据提取器 265 调用社交网络系统的模块以给所选择的用户呈现执行特定类型动作的机会并且监控所选择的用户是否执行动作。如果所选择的用户执行给予机会的动作,则数据提取器 265 记录动作类型的正面实例。如果所选择的用户不执行给予机会的动作,则数据提取器 265 记录动作类型的负面实例。动作的正面实例也被称为特定动作的成功,并且负面实例被称为特定动作的失败。

[0108] 数据提取器 265 可在动作数据存储器 260 存储所提取的数据。在动作数据存储器 260 中存储的信息描述由用户采取的与生成预测器模型或检测和评估预测器模型相关的各种动作。表 1 示出了给定类型动作的存储在动作数据存储器 260 中的实例数据值。

[0109] 表 1

[0110]

| 行 ID | 源 ID | 目标 ID | 目标类型 | 动作 ID | 标签 |
|------|------|-------|------|-------|----|
| 1001 | 100 | 200 | 用户 | 25 | 1 |
| 1002 | 100 | 201 | 群 | 35 | 0 |
| 1003 | 101 | 202 | 图像 | 65 | 1 |

[0111] 如表 1 所示,每行的数据具有表示唯一地识别行的行标志符的数据行 ID 值。源 ID 是唯一地识别动作的源的标志符,例如,执行或不执行动作的用户。目标 ID 代表表示对其执行动作的实体的标志符值。目标类型表示对应于目标的实体类型。目标能够是在社交网络系统中表示的任何类型的实体,例如,群、用户、图像、应用、评论、新鲜事项等。目标通常是源用户能够与其交互的实体。源用户执行对应于与目标实体的交互的动作。可以通过唯一标识符表示每种类型的动作。动作 ID 值存储与行表示的动作类型对应的标志符。例如,如果目标实体是图像并且动作可以是评论图像,或如果目标实体是用户,则动作可以对应于源用户发送请求以建立与目标用户的联系。在另一个实施方式中,如果动作在短时间段内由用户重复多次,则动作数据存储器 260 存储单个行。例如,如果用户在短时间段内检索多次描述另一用户的信息,或用户在短时间段内重复多次敲击针对实体的相同按钮,则为在相同源和目标之间的动作类型的所有这些实例存储单个行。在实施方式中,动作

数据存储器 260 存储表 I 的每行的计数,使得计数指示在过去的预定大小的时间间隔(例如,30 分钟的间隔)内通过相同用户针对相同目标实体执行的特定类型动作的实例数量。

[0112] 标签列指示行是对应于正面实例还是负面实例,例如,值 1 指示动作的正面实例,并且值 0 指示动作的负面实例。在表 I 中示出的实例数据中,具有 ID 100 的用户与动作的两个实例关联,一个实例对应于行 1001(其瞄准具有目标 ID 200 的实体),并且另一实例对应于行 1002(其瞄准具有目标 ID 201 的实体)。第三行对应于通过具有源 ID 101 的另一用户采取的动作实例。在一些实施方式中,标签能够占据除零和一之外的值,例如,标签可表示能够记录任何正整数值的计数器的值。

[0113] 基于在表 T1 中收集的事件,特征提取器 295 基于可获得的动作提取特征。特征的实例包括表征源用户和目标实体以及源和目标之间的关系的信息。例如,如果目标实体是用户,则特征可表示源用户在过去与目标用户通信的速度、目标用户访问描述目标用户的信息的速度、自从在源和目标用户之间的最后交互起的时间长度等。

[0114] 特征可基于与在社交网络系统 100 中表示的实体和在社交网络系统 100 中执行的动作相关的值表示为表达式。表示特征的这些表达式能够通过专家提供并且通过例如系统管理员的特权用户添加至系统。在实施方式中,特征被表示为在社交网络系统中记录的动作的函数,即,特征=函数(记录的_动作)(feature = function(logged_actions))。特征也能够是其他特征的函数,例如,基于其他特征或动作、或两者的组合的表达式。如作为表达式的特征的实例,如果目标是用户,并且观看_个人资料(view_profile)对应于源用户观看目标用户个人资料,观看_照片(view_photo)对应于源用户观看目标用户照片,并且观看_评论(view_comment)对应于源用户观看由目标用户张贴的评论,则称为观察的特征可以如下定义。

[0115] 观察=观看_个人资料+观看_照片+0.5×观看_评论

[0116] 在实施方式中,在以上等式中,术语的值,即,如果动作发生则观看_个人资料是 1,并且如果动作未发生则观看_个人资料是 0。在另一个实施方式中,每个术语的值可以是分数值,例如,基于描述特定动作的信息的分数,例如,在时间间隔或者与动作相关联的时间长度内(诸如用户在检索不同照片之前观察照片的时间长度)通过用户执行多次动作。

[0117] 在实施方式中,特征可以是基于相对于多个目标通过源用户执行的动作的合计值(aggregate value)。例如,特征可表示在给定时间间隔内通过源用户针对连接到源用户的所有其他用户执行的所有页面访问量的总和。另一个特征可表示用户观看通过连接到用户的所有其他用户张贴的图像的速度。特征可表示基于源用户相对于连接到源用户的所有其他用户的行为标准化(normalize)的通过源用户相对于目标用户执行的动作。例如,特征可表示使用源用户与连接至源用户的所有其他用户的交互平均量标准化的源用户与目标用户交互的频率。在特征存储器 250 中存储描述各种特征的元数据。

[0118] 在实施方式中,表 I 的数据表示形式可以是在动作数据存储器 260 中的数据库表。如他们在社交网络系统中发生的动作被添加至表 I 的数据表示形式。同样,添加基于动作类型更新为动作实例的特征。例如,如果特征基于用户观看照片的次数,则观看照片的用户的每个实例可引起重新评估特征。在实施方式中,特征表存储各种特征的值。例如,特征表可以具有列源 ID、目标 ID、目标类型、动作 ID、以及各种特征。

[0119] 在实施方式中,动作数据存储器 260 维持反馈表(其存储训练不同预测器模型需

要的训练数据)。每个预测器模型可与表示其中调用或使用预测器模型的背景 (context) 的调用位置相关联。可以通过在不同背景下的用户执行相同类型的动作。背景可与用户能够采取该动作的呈现给目标用户的特定信息相关联。例如,用户发送请求与另一用户连接的动作可在每个与用户界面类型关联的不同背景下执行。例如,当呈现用户潜在联系人列表时,用户可发送与另一用户连接的请求。替换地,当源用户观看描述目标用户的图像或其他信息时,可给源用户呈现与目标用户连接的机会。表 II 示出了包括源 ID、目标 ID、调用位置 (call site)、以及表示动作实例是成功还是失败的标签的列。在表 II 中呈现的信息可用于训练不同的预测器模型。

[0120] 表 II

[0121]

| 行 ID | 源 ID | 目标 ID | 标签 | 调用位置 |
|------|------|-------|----|------|
| 2001 | 100 | 200 | 1 | C1 |
| 2002 | 100 | 201 | 0 | C2 |
| 2003 | 101 | 202 | 1 | C3 |

[0122] 在实施方式中,各个特征的值可表示为与源和目标每个实例相关的名称值对。在另一个实施方式中,针对特定预测器模型生成的数据被表示为表 III,在表 III 中每个源和目标与和模型相关的各个特征相关联。表 III 还存储表示每个行的成功或失败的标签。

[0123] 表 III

[0124]

| 行 ID | 源 ID | 目标 ID | 标签 | 特征 F1 | 特征 F2 |
|------|------|-------|----|-------|-------|
| 2001 | 100 | 200 | 1 | 20 | 512 |
| 2002 | 100 | 201 | 0 | 20 | 630 |
| 2003 | 101 | 202 | 1 | 15 | 720 |

[0125] 以表 III 的表示形式存储的数据可针对与特定预测器相关的不同类型的使用分为不同部分。例如,数据的一部分可用于训练,另一部分可用于测试不同类型的机器学习模型,并且第三部分可用于评估不同的机器学习模型,例如,为比较不同类型的模型而选择给定调用位置的特定类型的模型。

[0126] 机器学习模块 275 实现能够针对每种给定动作类型尝试的各种类型的模型。通过机器学习模型 275 实现的不同类型机器学习模型的实例包括线性回归、逻辑回归、二次回归、幂回归、指数回归、提升树、权重决策树、以及其他合适类型的模型。每个模型可与特定系数值相关联。例如,线性回归可表示为 $w_1*f_1+w_2*f_2+w_3*f_3+\dots$,在此,每个特征 f_i 通过合适的系数 w_i 加权。

[0127] 模型评估器 280 基于例如使用记录 - 损失得分 (log-loss scoring) 为模型收集的数据评估每个预测器模型。模型评估器 280 可使用给定的模型预测结果并且将预测的

结果与观察的实际结果比较。在实施方式中,模型评估器 280 通过将由模型预测的结果与随机猜测进行比较来评估模型。在实施方式中,模型评估器 280 评估所有各种类型的模型的性能并且将它们性能彼此比较,从而选择最佳执行模型用于给定调用位置的预测器模型。如果可用于确定是否将现有模型替换为新模型,则模型评估器 280 也将所选择的模型与现有模型比较。如果比较两个不同的模型,则相对于相同数据组和比较的结果来评估它们。

[0128] 调度器 285 基于预定调度表调度用于再生成预测器模型的任务。每个预测器模型的针对特定动作的调度可以是不同的。在实施方式中,调度器 285 维持任务队列。每个任务识别特定动作类型的预测器模型以及对应于再生成预测器模型的时间的一个或多个时间戳值。在对应于时间戳的给定时间处,调度器选择任务并且调用合适模块来执行任务。预测器模型的再生成可对应于重新计算模型的各个系数。因此,可保留现有预测器模型,但是改变具体模型的系数。例如,如果现有预测器模型是线性回归模型,则新预测器模型也可以是线性回归模型(模型的各个项的系数利用新系数替换)。预测器模型的再生成可对应于重新生成新模型。例如,可生成新类型的模型以替换现有预测器模型。在一些情况下,如果新预测器模型不显著地优于现有模型或比现有模型差,则可保留现有预测器模型。

[0129] 整体处理

[0130] 图 3 示出了根据本发明的实施方式用于生成通过在社交网络系统中的各个模块调用的预测器功能的整体处理。社交网络系统 100 的各个模块可以通过预测器工厂 150 生成的预测器模型的消费者。例如,新鲜事生成器 235 可调用预测目标用户对观看经由新鲜事发送至目标用户的内容感兴趣的可能性的预测器模型。建议模块 140 可调用预测用户与向用户建议的群交互的可能性(例如,用户请求描述群的另一信息的可能性或用户成为群成员的可能性)的预测器模型。

[0131] 预测器工厂 150 接收请求 310 以生成由社交网络系统 100 其他模块使用的特定类型的用户动作的预测器模型。如上所述,请求可以是针对目标用户可能采取的某些类型动作(例如,描述另一用户的请求信息、建立与用户的联系、写入关于照片的评论、成为群成员等)的预测器模型。请求提供用于确定对应于动作的成功的确度和用于确定对应于动作的失败的度量。动作的成功对应于可用于确定是否给予用户执行给定类型的动作的机会、用户执行给定动作类型的实例的标准。动作的失败对应于可用于确定是否给予用户执行给定类型的动作的机会、用户不执行给定动作类型的实例的标准。能够基于从未执行动作直至移除执行动作的机会的事实确定动作的失败。当向用户呈现可以对应于执行不同动作的机会的一组新信息时可以发生此,从而重写对应先前动作的信息。替换地,能够基于来自用户的主动反馈确定动作的失败,例如,来自用户的指示用户不再想观看所呈现的信息、或用户不喜欢所呈现的信息的信号。因为用户向社交网络系统指示信息应当移除,所以用户不可能基于所呈现的信息执行动作,从而指示动作的失败。

[0132] 例如,如果动作对应于用户从呈现给用户的新鲜事请求额外的信息,则动作的成功可对应于用户点击新鲜事项以请求额外信息。类似地,对于这个实例,失败可对应于用户试图隐藏新鲜事项、用户指示不喜欢新鲜事项、或在向用户呈现新鲜事项之后用户不点击新鲜事项。

[0133] 在实施方式中,请求者也可规定针对其请求预测器模型的用户类型。请求的预测

器模型可瞄准具有某些人口统计特性的用户。例如,请求的预测器模型可确定在某个年龄群的女士成为在社交网络系统中表示的组织的成员的可能性或来自特定地理区域的男士对某些媒体内容感兴趣的可能性。

[0134] 模型生成器 290 生成 320 请求的预测器模型。本文中进一步描述模型生成器 290 如何生成 320 请求的预测器模型的细节,例如,在图 4 中示出的流程图。预测器模型可包括可执行代码,其能够与可执行代码相关联的某些参数或属性一起被调用。例如,预测器模型的可执行代码可以是机器学习模型,并且预测器模型的参数可以是由机器学习模型使用的系数。预测器模型在预测器存储器 120 中存储描述生成的预测器模型的信息。

[0135] 预测器工厂 150 还生成用于生成的预测器模型的句柄 (handle)。句柄可以是用于唯一识别预测器模型的任何值。句柄可以是用于调用预测器的函数、函数名称、或指针。预测器工厂可维持表映射处理程序 (table mapping handler) 以查找给予处理程序 (handler) 的预测器模型。句柄可以包括数字标志符、字符串、识别预测器模型的函数指针、或执行预测器模型的函数或方法的名称。句柄还包括识别对应于预测器模型的系数值的信息,例如,由机器学习技术使用的系数值。

[0136] 预测器工厂 150 将句柄返回 330 至请求者模块。请求者模块能够使用句柄以随后调用预测器模型。请求者模块可以提供预测器模型的句柄以及用于预测器模型具体调用的参数。例如,建议模块可通过将至生成的预测器模型的句柄以及识别具体用户和群的信息提供至预测器工厂来调用预测器模型,以预测用户请求描述群的信息的可能性。对于新请求可重复在图 3 中示出的处理的步骤,例如,对在社交网络系统的其他模块中使用的预测器模型的请求。

[0137] 图 4 示出了根据本发明的实施方式的生成 320 针对给定类型动作的预测器模型的处理。预测器工厂 150 响应于生成针对给定类型动作的预测器模型的请求执行这些步骤。预测器工厂 150 将通过数据提取器 265 生成的数据用作用于训练机器学习模型 275 的训练数据组。预测器工厂 150 训练不同类型的机器学习模型并且生成针对给定类型动作的不同预测器模型。预测器工厂 150 确定针对每种类型的模型的系数或参数值。

[0138] 当模型评估器 280 使用通过数据提取器 265 生成的数据的一部分作为评估 430 生成的不同模型的测试数据。模型评估器 280 可通过使用预测器模型在呈现给具体用户的给定情况下预测结果、并且然后实际上观察在给定情况下由用户执行的动作而评估给定模型。模型评估器 280 追踪预测结果匹配观察的结果的实例数量和预测器模型不能准确预测结果的实例数量。基于在预测针对给定类型动作的结果中它们的成功率给预测器模型计分。例如,针对给定类型动作预测正确结果占 70% 的预测器模型比针对相同类型动作预测正确结果占 40% 的预测器模型得分高。

[0139] 模型评估器 280 针对给定动作比较 440 不同类型的模型的性能以选择使用的模型。模型评估器 280 基于每种模型的性能选择模型。在实施方式中,模型评估器 280 选择执行最佳的预测器模型作为针对给定类型动作的预测器模型。预测器工厂 150 存储描述所选择的预测器模型的信息并且将至预测器模型的句柄返回至请求者。可随后使用预测器模型的句柄以调用用于预测针对给定类型动作的结果的预测器模型。

[0140] 图 5 是根据本发明的一个实施方式用于在社交网络系统中再生成模型并且更新它们的处理的流程图。如图 5 所示,调度器选择 510 针对特定类型的动作的预测器模型以再

生成并且重新评估预测器模型。数据提取器 265 提取 520 用于再生成针对该动作类型的预测器模型以及评估并且检测模型所需要的数据。模型生成器 290 生成 530 针对与所选择的预测器模型对应的动作类型的新预测器模型。模型评估器 280 评估 540 生成的新模型,例如,通过采样数据组并且执行预测器模型以预测动作并且将预测的结果与实际动作比较。模型评估器 280 将生成的 530 新预测器模型的性能与选择的 510 现有预测器模型比较 550。可以使用线上 A/B 测试比较来比较预测器模型。A/B 测试可使用在系统中比较的每个模型以观看模型对系统的影响。可使用终端用户的行为确定模型对系统的影响。如果确定 560 新预测器模型比现有模型执行得更好,则模型生成器 290 利用新预测器模型替换 570 现有模型。

[0141] 更新描述预测器存储器 120 中的预测器模型的数据以反应新生成的预测器模型。如果新预测器模型比现有模型执行得更差,或如果新预测器模型的性能方面相比于现有模型的改善低于阈值,则模型生成器 290 保持现有模型并且丢弃新预测器模型。调度器 285 可重新评估生成针对特定动作的预测器模型的速度。例如,如果在图 5 中示出的处理和每个迭代中生成的每个新预测器模型中观察的改善始终在先前的模型边界附近,则可以降低再生成预测器模型的速度。可通过调度器以不同的速度再生成针对每种类型动作的预测器模型。调度器可将再生成预测器模型的任务与时间戳的相关并且将该任务添加至队列。根据在队列中下一个任务的时间戳,调度器可选择 510 再生成的下一个预测器模型或等待一段时间并且随后再次检查队列。

[0142] 可替换应用

[0143] 说明书中描述的特征及优势并不是全部包括的,并且具体地,根据附图、说明书以及权利要求书,很多其他特征和优势对于本领域普通技术人员而言将变得显而易见。此外,应注意,在说明书中使用的语言主要是为了可读性和说明的目的而选择的,且并非旨在描绘或限定本发明的主题。

[0144] 例如,预测器模型可生成并且用于其他类型的在线系统并且不限于社交网络系统。例如,能够生成存储用户个人资料并且允许用户采取动作的在线系统和针对用户能够采取的各种动作的用户预测器。例如,在线系统可允许用户接收各种类型数据的馈送。预测器模型可开发为用于预测用户是否打算打开呈现给用户的馈送。预测器模型能够由在线系统使用以将呈现给用户的馈送进行排序,例如,可基于用户打算打开馈送或从馈送请求额外信息的可能性来给馈送排序。

[0145] 本发明实施方式的上述描述仅出于说明的目的而呈现,并且不旨在穷尽或者将本发明局限于所公开的精确形式。本领域技术人员可以理解,根据以上公开可以有多种修改和变形。

[0146] 本说明的一些部分就对信息操作的算法和符号表示方面描述了本发明的实施方式。这些算法描述和表示通常由数据处理领域中的技术人员使用以将他们工作的实质有效地传达给本领域其他技术人员。尽管在功能上、计算上或逻辑上描述了这些操作,但这些操作应被理解为由计算机程序或等效电路、微代码等来实现。而且,也已证明有时为了方便,这些操作是指作为模块的这些操作的配置,而不失一般性。所描述的操作及其相关模块可体现为软件、固件、硬件或者它们的任意组合。

[0147] 本文中所描述的任何步骤、操作或者处理可利用一个或多个硬件或软件模块单独

地或者结合其他设备执行或者实施。在一种实施方式中,软件模块可利用包括计算机可读介质(其包含计算机程序代码)的计算机程序产品来实施,该计算机程序代码能够由计算机处理器执行,以执行所述的任一或者所有步骤、操作或者处理。

[0148] 本发明的实施方式还可涉及用于执行本文中的操作的装置。该装置可被具体地构造为用于所需目的,和/或它可包括选择性激活或由存储在计算机中的计算机程序重新配置的通用计算设备。这种计算机程序可被存储在适用于存储电子指令并被耦接至计算机系统的总线的有形计算机可读存储介质或任何类型的介质中。而且,本说明书中提及的任何计算系统可包括单个处理器或者可以是采用多个处理器设计以提高计算能力的架构。

[0149] 最后,在说明书中使用的语言主要是为了可读性和指导性的目的而选择的,并且它可能没有被选择来描绘或限制发明的主题。因此,其旨在使本发明的范围不受具体实施方式限制,而是由基于其讨论的本申请的任何权利要求来限制。因此,本发明的实施方式的公开内容旨在是说明性的,而非限制在所附权利要求中所阐述的本发明的范围。

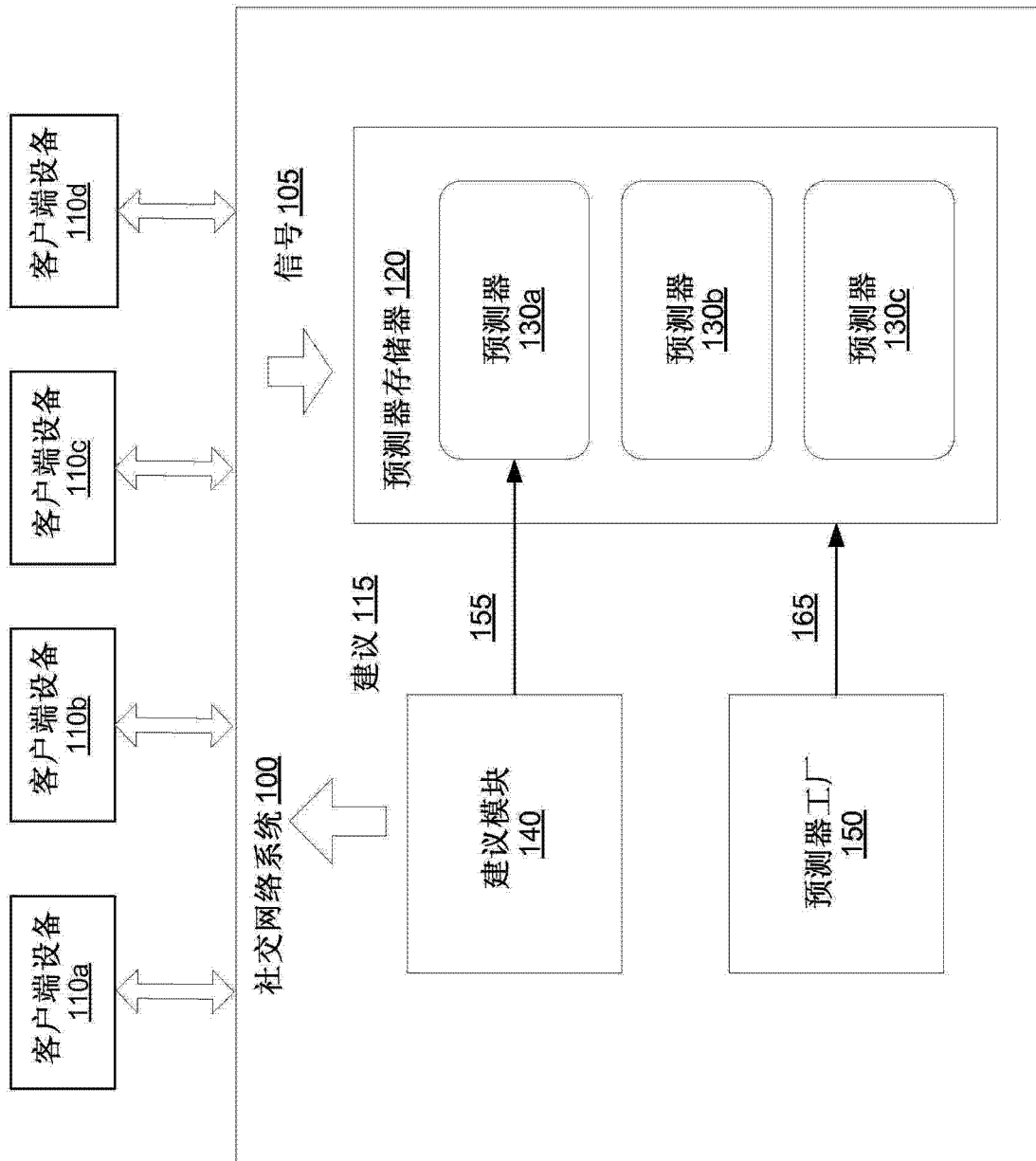


图 1A

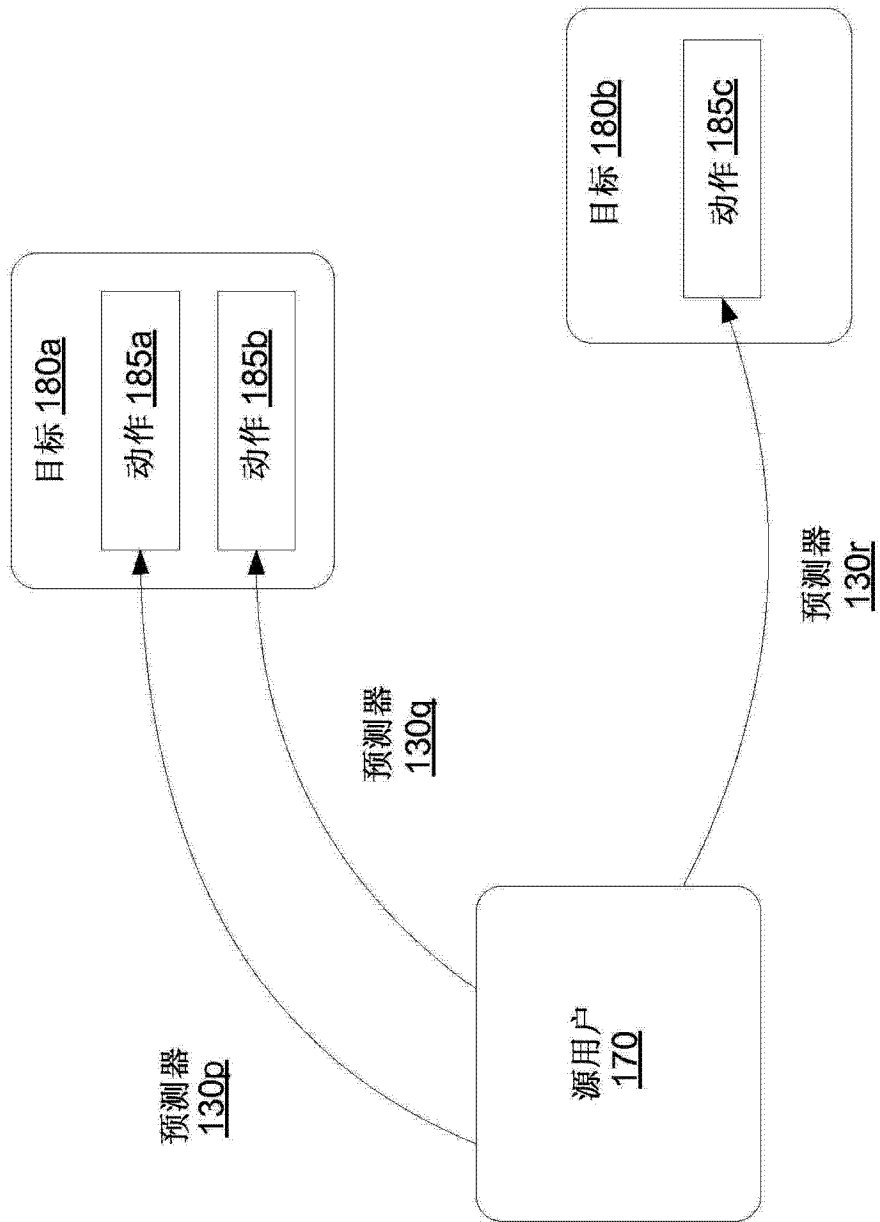


图 1B

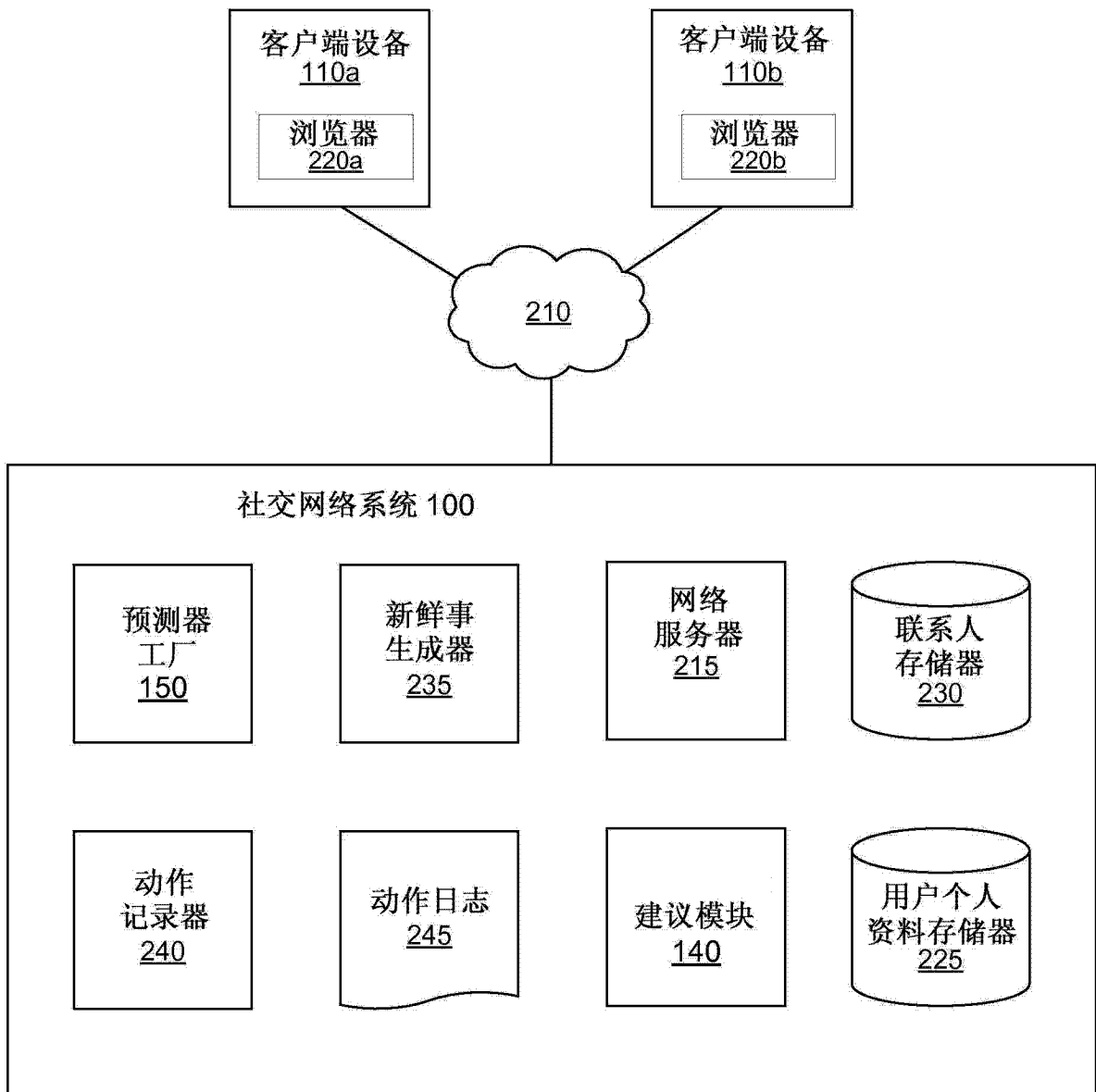


图 2A

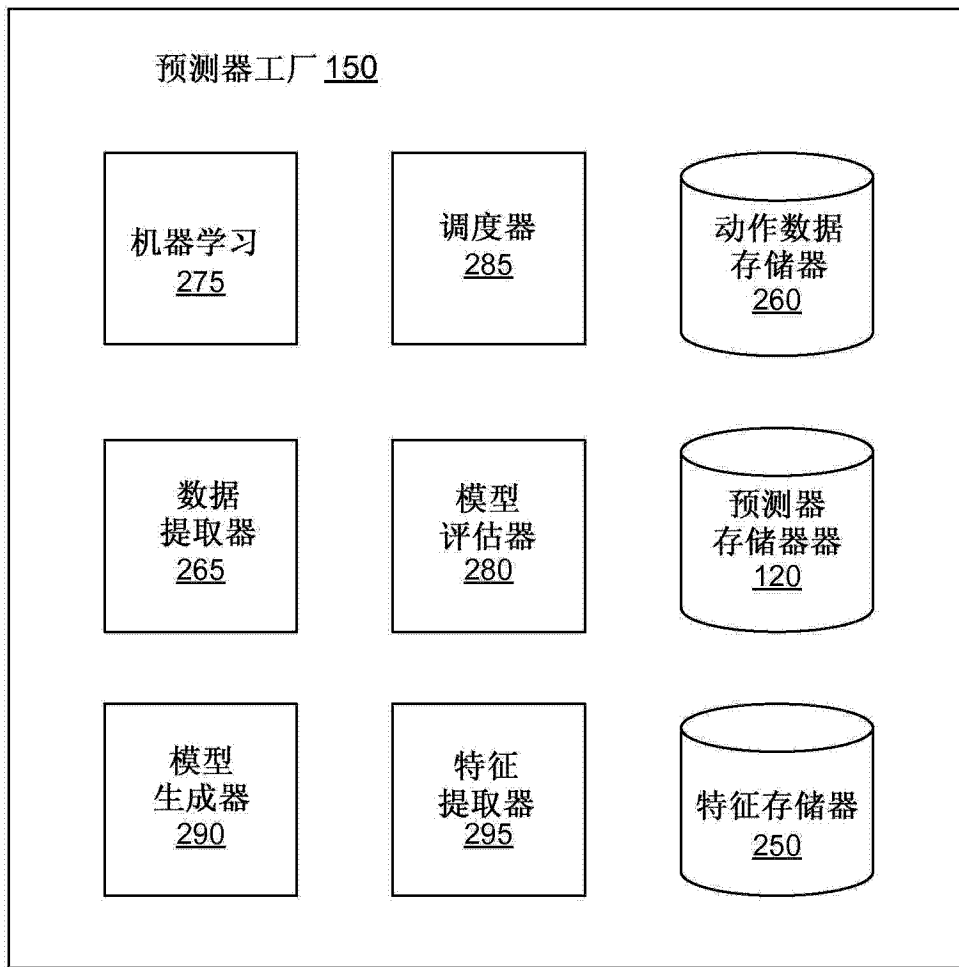


图 2B

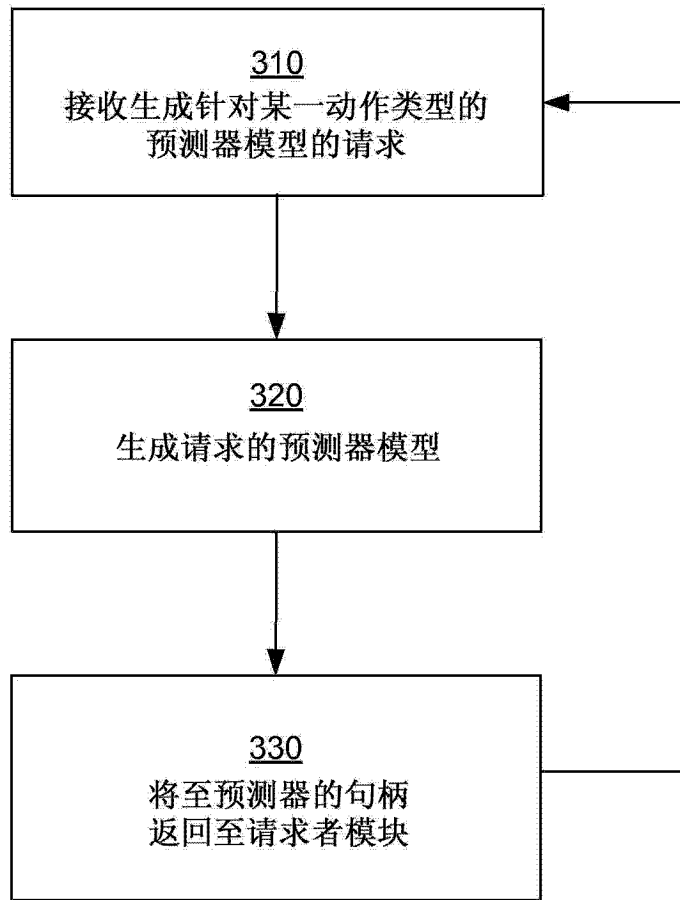


图 3

320

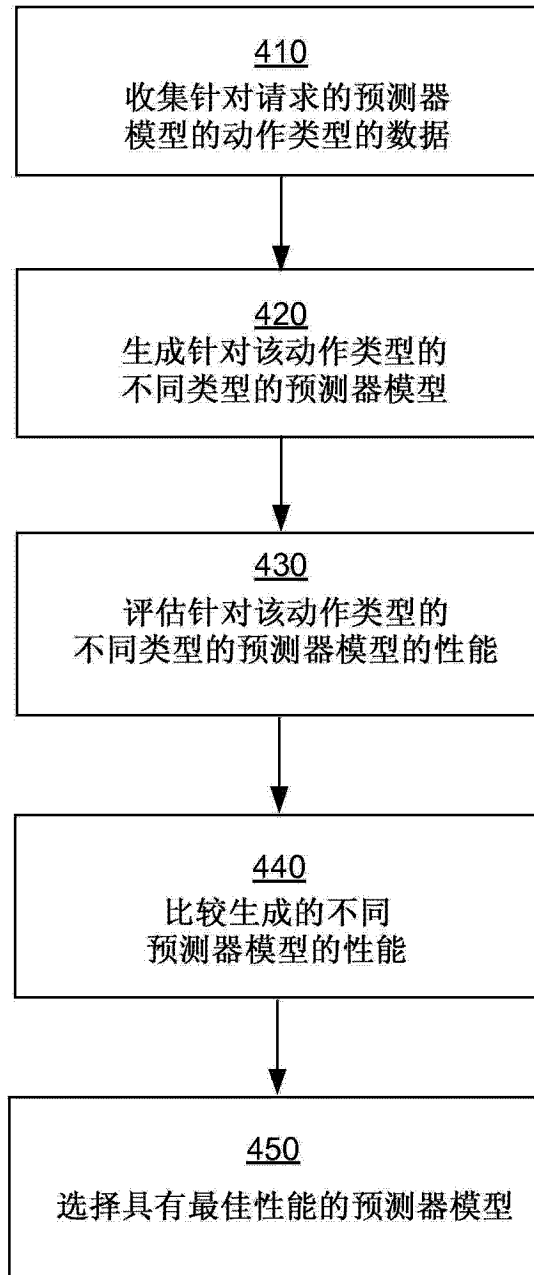


图 4

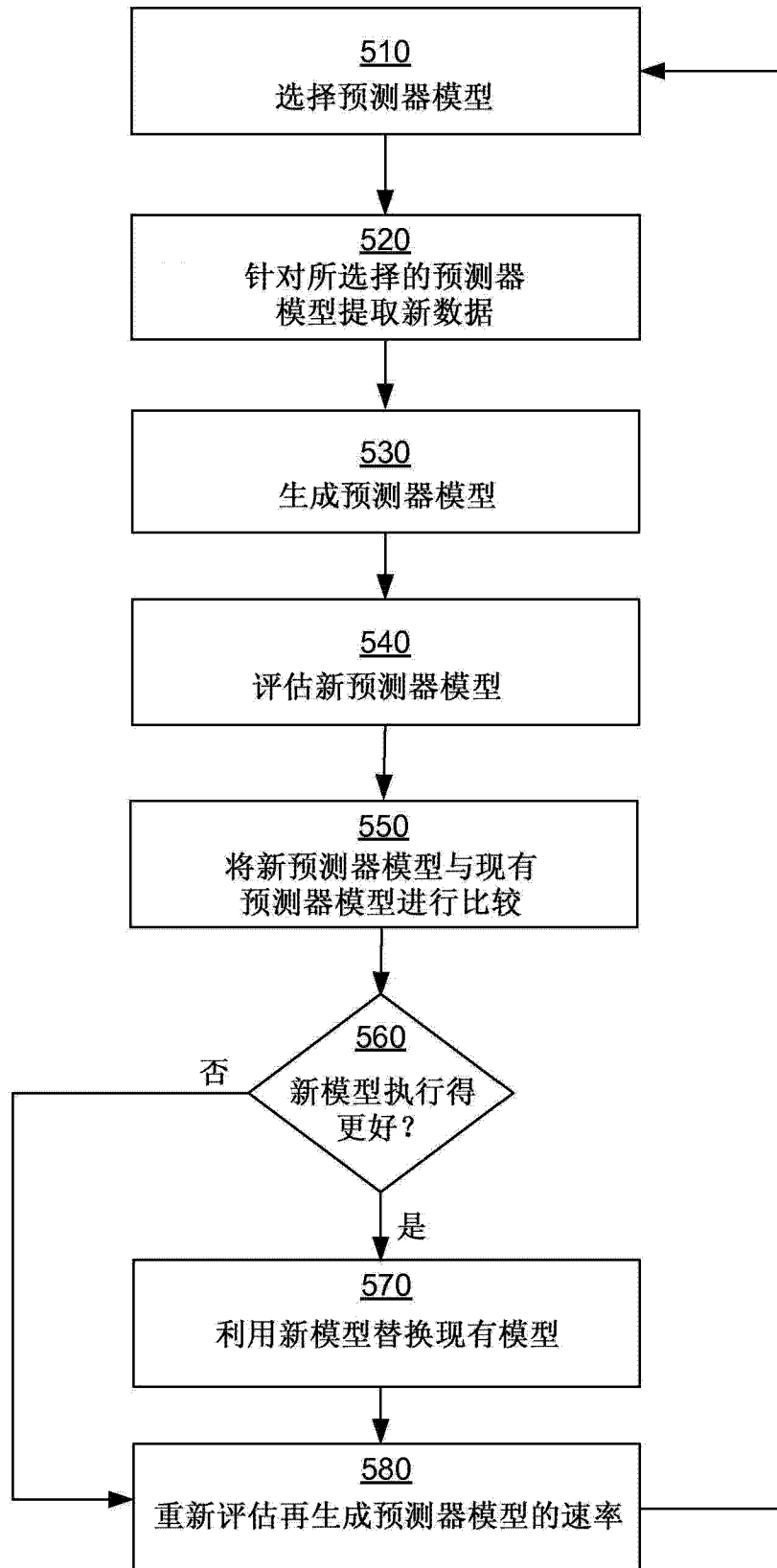


图 5