

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020年1月16日 (16.01.2020)



(10) 国际公布号
WO 2020/011068 A1

- (51) 国际专利分类号:
G06N 5/04 (2006.01)
- (21) 国际申请号: PCT/CN2019/094474
- (22) 国际申请日: 2019年7月3日 (03.07.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810751791.X 2018年7月10日 (10.07.2018) CN
- (71) 申请人: 第四范式(北京)技术有限公司(THE FOURTH PARADIGM (BEIJING) TECH CO LTD) [CN/CN]; 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。
- (72) 发明人: 王敏(WANG, Min); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。 李瀚(LI, Han); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。 乔胜传(QIAO,

Shengchuan); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。
陶雪军(TAO, Xuejun); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。 孙越(SUN, Yue); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。 唐继正(TANG, Jizheng); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。 徐昀(XU, Yun); 中国北京市海淀区上地东路35号颐泉汇大厦写字楼A座610室, Beijing 100085 (CN)。

(74) 代理人: 北京展翼知识产权代理事务所(特殊普通合伙)(ZYX INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区王庄路1号院清华同方科技大厦B座2层207, Beijing 100083 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

(54) Title: METHOD AND SYSTEM FOR EXECUTING MACHINE LEARNING PROCESS

(54) 发明名称: 用于执行机器学习过程的方法和系统

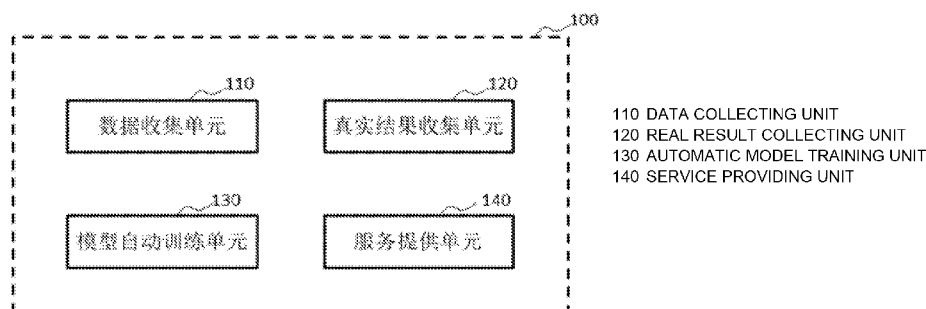


图 1

(57) Abstract: A method and a system (100) for executing a machine learning process, the system (100) comprising: a data collecting unit (110), configured to continuously collect prediction data; a real result collecting unit (120), configured to continuously collect real results of the prediction data; an automatic model training unit (130), configured to generate an updated training sample according to a configured model update scheme on the basis of the collected prediction data and the corresponding real results thereof, and using the updated training sample to continuously obtain updated machine learning models; and a service providing unit (140), configured to select from the machine learning models, according to a configured model application scheme, an online machine learning model for providing an online prediction service, and in response to a prediction service request including the prediction data, use the online machine learning model to provide a prediction result for the prediction data included in the prediction service request.

JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种用于执行机器学习过程的方法和系统(100), 所述系统(100)包括: 数据收集单元(110), 用于持续收集预测数据; 真实结果收集单元(120), 用于持续收集预测数据的真实结果; 模型自动训练单元(130), 用于按照配置的模型更新方案, 基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型; 以及服务提供单元(140), 用于按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型, 并响应于包括预测数据的预测服务请求, 利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

用于执行机器学习过程的方法和系统

技术领域

5 本申请总体说来涉及人工智能领域，更具体地讲，涉及一种用于执行机器学习过程的方法和系统。

背景技术

10 随着海量数据的出现，人工智能技术迅速发展，而机器学习是人工智能发展到一定阶段的必然产物，其致力于通过计算的手段，从大量数据中挖掘有价值的潜在信息。在计算机系统中，通过机器学习算法，可从历史数据中产生“模型”，也就是说，将历史数据提供给机器学习算法，就能基于这些历史数据进行建模来获得机器学习模型。

15 然而，在机器学习过程中，涉及数据、特征、算法、参数调优等诸多环节的处理，这都需要大量机器学习知识与经验；此外，如何在实践中利用训练好的模型来提供预测服务，也需要耗费大量的人工成本。上述各个环节都使得应用机器学习技术的门槛极高。即使已经存在了一些用于进行机器学习建模的平台产品，但现有的机器学习平台都仅注重于完成机器学习模型的调研，即，仅能完成如何基于积累的历史数据训练出一个机器学习模型，而无法有效地实现机器学习模型的后续生产和应用（例如，难以利用机器学习模型提供在线服务），换句话说，现有技术存在建模方案或成果与模型应用过程之间割裂严重的问题。

发明内容

20 本公开的示例性实施例在于提供一种用于执行机器学习过程的方法和系统，以至少解决现有技术存在的上述问题。

根据本公开的第一方面，提供一种用于执行机器学习过程的系统，其包括：

25 数据收集单元，用于持续收集预测数据；真实结果收集单元，用于持续收集预测数据的真实结果；模型自动训练单元，用于按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及服务提供单元，用于按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

30 根据本公开的第二方面，提供一种用于执行机器学习过程的计算机可读介质，其中，在所述计算机可读介质上记录有用于执行以下操作的计算机程序：持续收集预测数据；持续收集预测数据的真实结果；按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

35 根据本公开的第三方面，提供一种用于执行机器学习过程的计算装置，包括存储部件和处理器，其中，存储部件中存储有计算机可执行指令集合，当所述计算机可执行指令被所述处理器执行时，促使处理器执行以下操作：持续收集预测数据；持续收集预测数据的真实结果；按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

40 根据本公开的第四方面，提供一种用于执行机器学习过程的方法，包括：提供相互独立的第一操作入口和第二操作入口，其中，所述第一操作入口用于收集作为模型预测基础的行为数据，所述第二操作入口用于收集作为行为数据的真实结果的反馈数据；获取并保存通过所述第一操作入口收集的行为数据和通过所述第二操作入口收集的反馈数据；基于保存的所述行为数据和反馈数据，利用

至少一种模型算法来训练机器学习模型。

根据本公开的第五方面，提供一种用于执行机器学习过程的系统，包括：交互单元，用于提供相互独立的第一操作入口和第二操作入口，其中，所述第一操作入口用于收集作为模型预测基础的行为数据，所述第二操作入口用于收集作为行为数据的真实结果的反馈数据；数据收集单元，用于获取并保存通过所述第一操作入口收集的行为数据；真实结果收集单元，用于获取并保存通过所述第二操作入口收集的反馈数据；以及模型自动训练单元，用于基于保存的所述行为数据和反馈数据，利用至少一种模型算法来训练机器学习模型。

根据本公开的第六方面，提供一种用于用于执行机器学习过程的计算装置，包括本公开第一方面、本公开第五方面中任一项所述的系统；或者，

所述计算装置包括存储部件和处理器，其中，存储部件中存储有计算机可执行指令集合，当所述计算机可执行指令被所述处理器执行时，促使所述处理器执行根据本公开第四方面中任一项所述的方法。

根据本公开的第七方面，提供一种用于执行机器学习过程的计算机可读介质，其中，在所述计算机可读介质上记录有用于执行本公开第四方面中任一项所述方法的计算机程序。

根据本公开示例性实施例的用于执行机器学习过程的系统，使得数据收集、模型生产、模型应用等多个过程实现全流程循环作业，从而大大降低了机器学习技术落地的门槛和成本。

将在接下来的描述中部分阐述本公开总体构思另外的方面和/或优点，还有一部分通过描述将是清楚的，或者可以经过本公开总体构思的实施而得知。

附图说明

被结合在说明书中并构成说明书的一部分的附图示出了本公开的实施例，并且连同其说明一起用于解释本公开的原理。

图 1 示出根据本申请示例性实施例的用于执行机器学习过程的系统的框图；

图 2 示出根据本公开示例性实施例的用于执行机器学习过程的方法的流程图；

图 3 至图 8 示出根据本公开示例性实施例的用于执行机器学习过程的示例；

图 9 示出根据本公开另一实施例的用于执行机器学习过程的方法的流程图；

图 10 至图 17 示出根据本公开另一实施例的用于执行机器学习过程的示例；

图 18 示出根据本公开另一实施例的用于执行机器学习过程的系统的框图；

图 19 示出根据本公开另一实施例的用于执行机器学习过程的系统的框图；

图 20 示出根据本公开实施例的用于执行机器学习过程的计算装置的框图；

图 21 示出根据本公开另一实施例的用于执行机器学习过程的计算装置的框图。

具体实施方式

现在将详细参照本公开的实施例，所述实施例的示例在附图中示出，其中，相同的标号始终指的是相同的部件。以下将通过参照附图来说明所述实施例，以便解释本公开。在此需要说明的是，在本公开中出现的“并且/或者”、“和/或”均表示包含三种并列的情况。例如“包括 A 和/或 B”表示包括 A 和 B 中的至少一下，即包括如下三种并列的情况：(1) 包括 A；(2) 包括 B；(3) 包括 A 和 B。又例如“执行步骤一并且/或者步骤二”表示执行步骤一和步骤二中的至少一个，即表示如下三种并列的情况：(1) 执行步骤一；(2) 执行步骤二；(3) 执行步骤一和步骤二。

这里，为便于更好地理解本申请，首先对本申请中所涉及的部分知识进行解释。在机器学习领域，往往通过将历史数据提供给机器学习算法来训练“机器学习模型”，以确定构成机器学习模型的理想参数。训练好的机器学习模型可被应用于在面对新的待预测数据时提供针对相应预测目标的判断，即，预测结果。这里，可以看出，用于训练机器学习模型的历史数据（即，训练数据）作为机器学习的原材料，往往会导致产生不同效果的机器学习模型。不论是训练机器学习模型，还是利用训练好的机器学习模型进行预测，原始数据记录都需要转换为包括各种特征的机器学习样本。需要说明的是，在本申请中，除非另有明确说明，否则，本申请的示例性实施例对具体的机器学习算法并不进行特定限制。

<实施例一>

图 1 示出根据本申请示例性实施例的用于执行机器学习过程的系统 100 的框图。系统 100 包括数据收集单元 110、真实结果收集单元 120、模型自动训练单元 130 和服务提供单元 140。

具体说来，数据收集单元 110 可持续收集预测数据。这里，预测数据可以是用户（例如，用于推荐资

讯的资讯服务商)期望获得相关预测结果的数据,在这里,数据收集单元 110 可持续地从用户或者经由其他路径接收所述预测数据。例如,当用户想要获知期望向其客户(例如,终端消费者)推荐的资讯是否会被接受(即,是否会被消费者点击或阅读)的预测结果时,数据收集单元 110 可收集预测数据,即,关于期望推荐的资讯的属性信息数据。

5 作为示例,可在系统内部自动实现预测数据的收集,例如,用户可向服务提供单元 140 发送包括预测数据的预测服务请求,其中,预测数据可包括用户期望向消费者推荐的资讯和/或消费者的基本属性信息(例如,资讯主题、资讯展示位置、消费者标识符、性别、年龄、身高、体重、兴趣爱好等)。这里,系统 100 可提供预测数据自动回流功能,作为示例,该功能可被默认开启或根据用户的选择来开启,其中,该功能能够使服务提供单元 140 自动保存包括在预测服务请求中的预测数据,而数据收集单元 110 则可以持续地从服务提供单元 140 收集预测数据,例如,服务提供单元 140 可将预测数据打入特定集群(例如,Kafka 集群),再由数据收集单元 110 自动地从所述集群收集预测数据。可选地,预测数据也可以是用户(例如,银行)期望确定客户(例如,储户)是否存在欺诈行为的行为数据,但不限于此。为了便于描述,在下面的描述中,假定用户是向消费者推荐资讯的资讯服务商,而预测数据是包括用户期望向消费者推荐的资讯和/或相关消费者的基本属性信息的数据。

15 真实结果收集单元 120 可持续收集预测数据的真实结果。这里,预测数据的真实结果可以是指预测数据的真实标记(label),真实结果收集单元 120 可从用户或经由其他路径定期、批量或实时地收集预测数据的真实结果。例如,在用户想要预测推荐给消费者的资讯是否会被接受的示例中,预测数据的真实结果可指示消费者针对经过预测且实际推荐了的资讯的真实反馈。具体说来,针对预测数据的预测结果是服务提供单元 140 利用机器学习模型进行预测而得到的结果(例如,预测会点击阅读的结果表示为 1,预测不会点击阅读的结果表示为 0),在此基础上,可以进一步提供对于某条资讯而言消费者可能点击该资讯的概率,例如,0.9、0.85 或 0.76 等。真实结果收集单元 120 可持续地从用户收集与预测数据对应的反映消费者的真实反馈的真实结果,例如,在用户根据从服务提供单元 140 接收到的针对预测数据的预测结果(例如,消费者可能点击各个资讯的概率)而向消费者推荐若干资讯后,消费者对所述若干资讯中的至少一条资讯进行了点击浏览,而没有理会所述若干资讯中的其余资讯,那么与所述至少一条资讯对应的预测数据的真实结果均可被赋予 1,而与所述其余资讯对应的预测数据的真实结果均可被赋予 0。

20 模型自动训练单元 130 可按照配置的模型更新方案,基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型。

30 这里,模型更新方案可以是用于完整训练模型的方案,也可以是用于在已有模型的基础上进行增量学习的方案。相应地,随着不断收集预测数据及其真实结果,可按照模型更新方案持续地产生新的模型。

45 作为示例,在模型自动训练单元 130 持续得到更新的机器学习模型之前,系统 100 中已经存在一个初始机器学习模型,该初始机器学习模型可以由系统 100(例如,模型自动训练单元 130)先前利用模型训练方案训练出的机器学习模型,也可以是由其它系统产生的并被导入到系统 100 中的机器学习模型。利用该初始机器学习模型,系统 100(例如,服务提供单元 140)可在初始阶段利用其提供在线预测服务。这里,作为示例,模型自动训练单元 130 可在训练出初始机器学习模型的模型训练方案基础上生成配置的模型更新方案,也就是说模型更新方案在数据、特征、算法和/或参数等方面的处理可与所述模型训练方案一致。在此基础上,模型自动训练单元 130 可按照配置的模型更新方案利用更新训练样本来持续地得到更新的机器学习模型。

40 这里,更新训练样本是具有真实结论的数据样本,换句话说,更新训练样本可以是按照模型更新方案中限定的与特征生成有关的过程,通过对收集的预测数据进行特征抽取而得到特征后结合该预测数据对应的真实结果(作为标记,即,label)所产生的样本。具体地讲,由数据收集单元 110 收集的预测数据可例如以数据表的形式被存储在本地存储介质或具有数据存储功能的云计算平台(包括但不限于公有云 and 私有云)中。通常,数据表的一行可对应于一条数据记录,数据表的一列可对应于一个属性字段。数据表中的每条数据记录可包括一个或多个属性信息(即,属性字段)。属性字段可用于形成特征,而特征是构成训练样本的组成部分,这里,特征可以是属性字段本身或属性字段的局部、属性字段的组合、或属性字段经过处理(或运算)而得到的结果。不同特征之间还可以进一步结合。例如,数据表中的一行预测数据经过特征抽取后可得到至少一个特征,而所得到的所述至少一个特征以及该行预测数据的对应的真实结

果就组成了一个更新训练样本。优选地，更新训练样本可以是通过对收集的预测数据进行特征抽取和特征组合而得到各种特征后结合该预测数据的真实结果所产生的样本。

在产生了更新训练样本后，模型自动训练单元 130 可进一步按照配置的模型更新方案中限定的关于模型训练步骤方面（例如，模型算法、参数调优等）的设置，利用更新训练样本持续得到更新的机器学习模型。如上所述，配置的模型更新方案可以是模型自动训练单元 130 在训练出初始机器学习模型的模型训练方案基础上生成的，也可以是任何用于不断训练出机器学习模型方案，这里的模型更新方案旨在强调该方案可用于较为自动地不断生成模型，并不限制模型生成的方式为全部重新训练还是增量学习训练。

此外，为了更好地自动执行模型更新，模型更新方案还可包括以下项之中的至少一项：用于从预测数据中选取用于生成更新训练样本的预测数据的数据选取规则、用于由模型自动训练单元更新机器学习模型的模型更新周期、模型存储位置、更新资源自动配置方式。

具体地讲，对于数据选取规则而言，由数据收集单元 110 所收集的预测数据可能包含非常多的预测数据，而数据选取规则可规定具体选取由数据收集单元 110 所收集的预测数据中的哪些预测数据来产生更新训练样本，例如，选取全部数据、按时间范围选取数据（例如前 3 天数据）、按存储位置的范围选取数据（例如第 20 切片数据~第 200 切片数据）等规则。

对于模型更新周期而言，为了使机器学习模型能够针对接收的预测数据提供更加准确的预测结果，模型自动训练单元 130 可按照一定的模型更新周期来更新机器学习模型（即，产生新的机器学习模型），该模型更新周期可由用户预先配置好，也可以按照一定规则根据具体情况而实时修改。

对于模型存储位置而言，由于机器学习模型的不断更新，会得到多个机器学习模型，为了使服务提供单元 140 能够从多个机器学习模型中选择用于提供在线预测服务的线上机器学习模型，模型自动训练单元 130 需要确定存储持续得到的更新的机器学习模型的位置，例如，机器学习模型可被存储在系统 100 内置的模型中心中，这还可使得用户能够查看模型相关的解释和报告。

对于更新资源自动配置方式而言，模型自动训练单元 130 在得到更新的机器学习模型的过程中需要知道按照何种方式利用系统资源（例如，CPU、总线、带宽、内存等资源），这里，模型自动训练单元 130 可以按照数据量结合规则来配置资源，但本公开不限于此。

模型训练方案可以由模型自动训练单元 130 在利用自动机器学习技术训练出初始机器学习模型的过程中所确定的方案，也可以是通过其它方式提前确定的任意适合的模型训练方案。下面将详细描述模型训练单元 130 如何在训练初始机器学习模型的过程中确定模型训练方案的过程。

根据本公开的示例性实施例，初始机器学习模型可指示在模型调研阶段根据自动机器学习技术训练出的模型，作为示例，其可被用于在最初阶段针对预测数据提供预测服务，将训练所述初始机器学习模型时所基于的数据记录称为历史数据，为此，数据收集单元 110 除了可以收集预测数据之外，还可以收集历史数据，相应地，真实结果收集单元 120 除了可以收集预测数据的真实结果之外，还可以收集历史数据的真实结果，其中，历史数据是历史上累积的已经具有真实结果的数据。这里，作为示例，历史数据记录本身可能包括标记字段（即，真实结果），但为了与本公开示例性实施例的系统相一致，可首先对历史数据记录表做切分，以得到历史数据（不包含标记字段）和标记字段两部分，并将历史数据和真实结果分别导入数据收集单元 110 和真实结果收集单元 120。

由于建模过程是机器学习门槛较高的环节，在建模过程中，数据处理、特征工程、算法选择、调参等都需要大量机器学习知识与经验，因此将整个建模过程自动地串联起来实现自动机器学习过程的难度则非常大并且极少能够实现，本公开旨在通过模型自动训练单元 130 来实现自动机器学习过程。下面将对此进行详细描述。

模型自动训练单元 130 可根据自动机器学习技术，基于收集的历史数据及其对应的真实结果生成初始训练样本并利用初始训练样本训练出初始机器学习模型。初始训练样本可以是根据自动机器学习技术通过对收集的历史数据进行特征抽取而得到特征后结合该历史数据对应的真实结果而产生的样本，这里，作为示例，可根据历史数据的各字段的数据类型和/或用于训练模型的算法将各字段自动声明为离散特征或连续特征。优选地，初始训练样本可以是根据自动机器学习技术通过对收集的历史数据进行特征抽取和特征组合而得到各种特征后结合该历史数据的真实结果而产生的样本。

自动机器学习技术可涉及以下项中的至少一项：用于将历史数据拆分为训练数据和验证数据的自动数据拆分、用于按照数据字段类型和/或数据分布对训练数据和验证数据进行特征抽取的自动特征生成、用于

按照特征重要性来确定组合特征的自动特征组合、用于按照预设调参方式进行参数调优的自动调参、用于根据预测效果确定由一个或多个训练出的模型来提供预测结果的模型自动选择、以及根据预设规则对资源的自动配置。

具体说来，可按照预设拆分规则来自动地将历史数据拆分为训练数据和验证数据，这两部分数据集将经历同样的特征处理过程。相应地，可按照数据字段的类型（例如，数值型、类别型等）和/或数据的分布特性，进行自动的特征生成过程，例如，将数据集中某些字段自动声明为离散特征或连续特征，或者进行特定的数值运算（如，取对数运算等等）。优选地，可根据关于如何确定特征重要性的算法来确定将哪些特征进行组合，例如，可按照规则构建多种候选的特征组合方式，针对每种候选特征组合方式，确定该组合方式下产生的组合特征的相对重要性（例如，基于该特征在模型中的预测作用来衡量该特征的重要性），并将重要性较高的特征组合确定为最终采用的特征组合方式。模型自动训练单元 130 还可采用自动调参的方式（例如，自动选取最优解等）在模型训练的过程中进行自动调优。这里，模型自动训练单元 130 可针对历史数据的特点选择指定的模型算法来进行训练，为此，作为可选方式，模型自动训练单元 130 可指定一种或多种模型算法来同时训练多个模型，并可根据这些模型的效果来确定由哪些模型最终提供预测服务，或者将各个模型提供的预测结果进行加权来作为最终结果。此外，在进行以上自动机器学习的过程中，还需要根据预设规则对系统资源（例如，CPU、带宽、内存等）进行自动配置，例如，可以按照数据量结合规则来配置系统资源，但本公开不限于此。经过以上的自动机器学习过程，与所选择的效果最佳的机器学习模型相应的预置算法及其经过自动调参所确定的相应的参数被确定为模型训练方案。

如上所述，根据自动机器学习技术，基于历史数据样本得到的初始机器学习模型可直接用来提供在线服务，相应的方案可作为模型训练方案。优选地，上述初始机器学习模型可不直接用于提供在线服务，而是由按照所述模型训练方案重新训练出来的模型作为最初提供在线服务的模型，在重新训练的过程中，训练数据进一步涵盖之前的验证数据，由于增大了训练数据量，上述方式产生的模型可得到更好的预测效果。

以上描述了利用自动机器学习技术得到模型训练方案的过程，在该模型训练方案的基础上，模型自动训练单元 130 可进一步结合数据选取、更新频率、模型存储位置、资源配置等方面形成模型更新方案。

相应地，模型自动训练单元 130 可按照配置的模型更新方案来利用以上所描述的更新训练样本持续得到更新的机器学习模型。

作为示例，在得到更新的机器学习模型的过程中，模型自动训练单元 130 可被设计为仅按照增量学习方式，利用以上所描述的更新训练样本对机器学习模型进行更新。作为另一示例，模型自动训练单元 130 可被设计为仅按照全量学习方式，根据模型训练方案利用以上所描述的更新训练样本重新训练出机器学习模型，以作为更新的机器学习模型。可选地，模型自动训练单元 130 可被设计为根据当前提供在线服务的机器学习模型的效果来确定到底是使用增量学习方式还是使用全量学习方式来得到更新的机器学习模型，例如，如果当前提供在线服务的模型的效果变得非常差（例如，AUC 小于预定阈值），则模型自动训练单元 130 可按照全量学习方式来重新训练机器学习模型。如果当前提供在线服务的模型的效果尚可（例如，AUC 大于或等于预定阈值），则模型自动训练单元 130 可按照增量学习方式来更新机器学习模型。可选地，模型自动训练单元 130 可被设计为根据用户设置来确定是使用增量学习方式还是使用全量学习方式来得到更新的机器学习模型。

此外，由于数据收集单元 110 可持续地收集预测数据并且真实结果收集单元 120 可持续地收集预测数据的真实结果，因此模型自动训练单元 130 可按照模型更新方案中包括的模型更新周期根据上述过程持续得到不断更新的机器学习模型，并将初始机器学习模型和持续得到的更新的机器学习模型存储在模型更新方案中规定的模型存储位置处。

除了通过自动机器学习过程来得到机器学习模型存在一定难度之外，机器学习模型在应用过程中同样困难重重，这是因为机器学习模型在提供预测服务过程中需要解决数据在线拼接、在线特征工程（即，自动特征抽取，还可包括自动特征组合）、模型选取、资源配置、服务监控等问题，然而目前机器学习模型的应用过程普遍存在产品化程度较低的问题。然而本公开却可以通过系统 100 中的各个单元有效地解决这些问题。

服务提供单元 140 可按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括

在预测服务请求中的预测数据提供预测结果。

具体地讲，如上所述，模型自动训练单元 130 可将训练出的机器学习模型不断存储在模型更新方案中规定的模型存储位置处，其中，训练出的机器学习模型可包括初始机器学习模型和持续得到的更新的机器学习模型。在此情况下，服务提供单元 140 可按照配置的模型应用方案从存储的机器学习模型中选择用于提供在线预测服务的线上机器学习模型，并向外提供关于预测服务的应用编程接口（API）。在这种情况下，用户可以通过所述 API 请求针对相应预测数据的预测服务（即，请求系统 100 针对预测数据，提供关于预测目标的预测结果）。

作为示例，模型应用方案可包括用于从机器学习模型之中选择线上机器学习模型的模型选取规则和/或应用资源自动配置方式。根据预设的模型选取规则，可自动选取上线模型，例如，可选择具有最高 AUC 的机器学习模型或最新产生的机器学习模型作为线上机器学习模型，但本公开不限于此，例如，除了上述自动的模型选取规则之外，还可以采用手动方式来选择上线模型。或者，将自动和手动方式相互结合，即，既设置用于选取线上机器学习模型的规则，同时提供人为确认或评估的交互方式。应用资源自动配置方式可以是在应用所选择的线上机器学习模型时如何对系统资源进行配置，例如，可按照数据量结合规则来配置系统资源，也根据请求的流量动态设置资源，但本公开不限于此。

此外，为了使所选择的机器学习模型能够自动地对来自用户的预测数据提供预测结果，需要使得服务提供单元 140 能够自动地对包括在预测服务请求中的预测数据进行特征抽取来得到适合于线上机器学习模型的预测样本。为此，模型自动训练单元 130 将训练出的机器学习模型的相应的特征抽取过程信息包括在机器学习模型所对应的模型文件的元数据中，换句话说，模型自动训练单元 130 得到的机器学习模型所对应的模型文件的元数据中可包括相应的特征抽取过程信息。作为示例，特征抽取过程信息可包括关于数据字段的处理、从数据字段生成单位特征的处理和/或对单位特征进行进一步运算或组合的处理，应注意，本公开的示例性实施例对特征抽取过程所涵盖的处理细节不进行任何限制，只要其有助于将接收的预测数据转换为能够输入机器学习模型的预测样本。因此，在服务提供单元 140 所选择的线上机器学习模型被用于提供预测结果之后，服务提供单元 140 可利用线上机器学习模型所对应的文件中的特征抽取过程信息自动地对预测服务请求中的预测数据进行特征抽取以得到预测样本，从而可实现在线特征工程，此后，服务提供单元 140 可利用线上机器学习模型针对预测样本提供预测结果。此外，在服务提供单元 140 利用线上机器学习模型提供在线预测服务时，用户可监控模型应用服务的状态和日志等。

举例来说，在用户是向消费者推荐资讯的服务商的情况下，当该用户将包括预测数据（即，多条资讯以及消费者的基本属性信息（例如资讯主题、资讯展示位置、消费者标识符、性别、年龄和兴趣爱好等））的预测服务请求发送到服务提供单元 140 时，服务提供单元 140 可自动地保存预测服务请求中的预测数据以供数据收集单元 110 收集预测数据，并且服务提供单元 140 利用线上机器学习模型所对应的文件中的特征抽取信息自动地对这些预测数据进行特征抽取来产生预测样本，此后，服务提供单元 140 将预测样本输入在线机器学习模型并最终获得预测结果，例如，该预测结果可以是对于某条资讯而言消费者可能点击该资讯的概率，例如，0.9、0.85、0.76 等。此后，服务提供单元 140 将该预测结果提供给用户，而用户可根据接收的预测结果（例如，消费者可能点击各个资讯的概率）将消费者最可能点击的若干条资讯发送给消费者。在这种情况下，如果消费者点击了由提供资讯的用户所提供的某条资讯，则该用户可将与该条资讯对应的预测数据的真实结果赋值为 1，如果消费者没有点击由用户提供的某条资讯，则该用户可将与该条资讯对应的预测数据的真实结果赋值为 0。在此基础上，真实结果收集单元 120 可从用户持续收集预测数据的真实结果。通过上述过程可以实现预测数据及其真实结果的自动回流，并且可使得模型自动训练单元 130 能够利用回流的预测数据及其真实结果进一步产生更新的机器学习模型，从而实现自动机器学习过程。

此外，在利用线上机器学习模型提供在线预测服务时，服务提供单元 140 还会自动地保存针对预测数据的预测结果，并根据由真实结果收集单元 120 收集的与预测数据相应的真实结果自动统计各项指标（包括正样本率、预测 AUC 和其他业务指标等）来进一步评估模型上线效果。

通过以上描述可以看出，系统 100 有效地消除了现有技术中机器学习过程与应用过程严重割裂的问题以及由此造成的数据回流、模型效果评估、业务指标统计无法串联的问题，使得企业不需要针对不同模型单独地定制开发，可使得模型可在企业内部被复用和沉淀。

为了使本领域技术人员能够更好地理解本公开，下面将参照附图 2 中示出的用于执行机器学习过程的方法的流程图以及图 3 至图 8 所示出的图形用户界面来描述本公开。应注意，图 2 中示出的

步骤顺序仅用于对本公开进行示例性说明，并不在于限制各步骤的发生顺序，实际上，本公开示例性实施的各步骤并非一定具有时间上的先后关系，特别是在处理不断流入的数据的情况下，完全可以按照其它任何合适的步骤顺序进行或者在时间上完全或部分重叠地执行。

如图 2 中所示，在步骤 S201，数据收集单元 110 收集历史数据，其中，历史数据是历史上累积的数据，5 这些数据具有对应的真实结果，在下面的描述中，假设历史数据来自于向消费者推荐资讯的资讯服务商，系统 100 旨在训练用于预测消费者是否会接受被推荐的资讯的机器学习模型，并将适当的机器学习模型用于提供关于上述预测结果的预测服务。这里，历史数据可以以数据表的形式被导入系统 100，数据表的一行代表一条历史数据记录，每条数据记录都包括资讯以及消费者的基本属性信息（例如，资讯主题、资讯展示位置、消费者标识符、性别、年龄、身高、体重、兴趣爱好等）。

具体地讲，如图 3 中所示的图形用户界面，为了导入历史数据，用户可点击该图形用户界面的左上方的与行为数据对应的“进入”按钮 401，然后进入图 4 的图形用户界面。在图 3 所示的界面中，为了使得机器学习过程更加容易被理解，将“行为数据”这一表述用于提示“历史数据”和“预测数据”的导入，将“反馈数据”这一表述用于提示真实结果数据的导入。

在图 4 中，可通过三种方式来导入历史数据，即，将本地存储的历史数据上传系统、通过分布15 式文件系统（Hadoop Distributed File System, HDFS）导入历史数据、通过 Kafka 实时流入历史数据（这里，应注意，尽管页面中示出了上述三种导入方式，但可根据具体应用情景而只启用其中的一种或多种）。例如，当选择了图 4 中所示的“本地上传数据”按钮之后，用户可从本地存储的历史数据中选择想要上传系统的历史数据。图 5 示出了已经被用户选中的将被上传系统的历史数据，其中，历史数据的表共有 37000 行和 37000 列。此外，为了使历史数据与其真实结果之间能够一一对应，20 还需要对历史数据的唯一 ID、数据列的类型等信息进行配置。当用户点击图 5 中的“下一步”之后可将该历史数据上传到系统中，即，被数据收集单元 110 收集。

在步骤 S202，真实结果收集单元 120 可收集历史数据的真实结果，其中，历史数据的真实结果是指历史数据记录所对应的真实标记字段。在上述示例中，真实结果是指反映了消费者是否接受了相关资讯的真实结果，例如，是否点击阅读了这些历史上推荐的资讯。作为示例，如果消费者对所述若干条历史资讯中的至少一条历史资讯进行了点击浏览，而没有理会所述若干条历史资讯中的其余历史资讯，那么与所述至少一条历史资讯对应的真实结果均可被赋予标志 1，而与所述其余历史资讯对应的真实结果均可被赋予标志 0。

具体地讲，如图 3 所示，当用户点击了图 3 的图像用户界面中的右上方的与“反馈数据”对应的“进入”按钮 402 时，可进入图 6 的图形用户界面。如图 6 中所示，同样可以通过三种方式来导入真实结果，30 即，将本地存储的真实结果上传系统、通过 HDFS 导入真实结果、通过 Kafka 实时流入真实结果（这里，应注意，尽管页面中示出了上述三种导入方式，但可根据具体应用情景而只启用其中的一种或多种）。由于真实结果的导入方式与历史数据的导入方式类似，因此，此处不再进行详细描述。此外，为了使历史数据与其真实结果之间能够一一对应，同样需要对与历史数据相应的真实结果的唯一 ID、数据列的类型等信息进行配置。通过上述过程，用户可将真实结果上传到系统中进而被真实结果收集单元 120 收集。

在步骤 S203，模型自动训练单元 130 可根据自动机器学习技术，基于收集的历史数据及其对应的真实结果生成初始训练样本并利用初始训练样本训练出初始机器学习模型。初始训练样本可以是根据自动机器学习技术通过对收集的历史数据进行特征抽取而得到特征后结合该历史数据对应的真实结果而产生的样本，优选地，初始训练样本可以是根据自动机器学习技术通过对收集的历史数据进行特征抽取和特征组合而得到各种特征后结合该历史数据的真实结果而产生的样本。应注意，35 在进行特征抽取之前，需要利用历史数据的 ID 和真实结果的 ID 对历史数据和真实结果进行拼接。由于以上已经详细地描述了产生初始训练样本的详细过程，因此此处不再进行赘述。以下将参照图 7 和图 8 对自动机器学习技术进行详细描述。

具体地讲，图 7 中间部分示出的有向无环图（DAG 图）示出了 6 个节点：“反馈数据”节点、“行为数据”节点、“数据拆分”节点、“特征工程”节点、“LR（逻辑回归）算法”节点、“GBDT（梯度提升决策树）算法”节点、“HE-TreeNet（高维离散嵌入式的树网络）算法”节点和“NN（神经网络）算法”节点。应注意，图 7 示出了 4 种具体预置算法，但是这仅是示例性说明，本公开并不对预置算法的数量和具体算法进行限制，此外，图 7 中的 DAG 图并没有明确地示出模型自动训练单元 130 所使用的自动机器学习技术所涉及的全部内容。

参见图 7，经由 DAG 图中的“数据拆分”节点处的相应配置，模型自动训练单元 130 可将历史数据拆分为训练数据和验证数据。此后，经由 DAG 图中的“特征工程”节点处的相应配置，模型自动训练单元 130 可对拆分出的训练数据/验证数据进行自动特征生成来抽取至少一个特征，优选地，模型自动训练单元 130 还可在自动特征生成之后进行自动特征组合以获得包括组合特征的各种特征。对应于 DAG 图中的最下层的四个节点（即“LR 算法”节点、“GBDT 算法”节点、“HE-TreeNet 算法”节点和“NN 算法”节点）处，模型自动训练单元 130 利用经过特征工程之后形成的训练样本/验证样本对这四个预置算法分别进行训练（结合自动调参），进而训练出四个机器学习模型，在该示例中，模型自动训练单元 130 根据自动机器学习技术，训练出四个机器学习模型，但本公开的示例性实施例不限于此。此外，在该示例中，模型自动训练单元 130 可以根据模型效果等因素从四个机器学习模型中选取一个或多个机器学习模型作为初始机器学习模型，其中，如果多个机器学习模型被选取，则可综合这些机器学习模型的预测结果来得到向外提供的预测结果。此外，在进行以上过程中，还需要根据预设规则对系统资源（例如，CPU、带宽、内存等）进行自动配置，例如，可以按照数据量结合规则来配置系统资源，但本公开不限于此。

作为示例，假设模型自动训练单元 130 将 GBDT 算法对应的方案确定为模型训练方案，例如，图 8 中所示的 GBDT 模型训练方案。在这种情况下，模型自动训练单元 130 可按照所述模型训练方案，基于未进行自动数据拆分的历史数据整体（即，包括训练集和验证集两者）来再次训练出 GBDT 模型，以用于面向最初的预测数据在线提供预测结果。可选地，模型自动训练单元 130 可不重新产生初始提供预测服务的机器学习模型，而将在自动机器学习过程中确定模型训练方案的同时所产生的初始机器学习模型确定为初始的在线机器学习模型。在这种情况下，在步骤 S203 获得的初始机器学习模型可被用作最初的用于提供在线预测服务的线上机器学习模型。

虽然目前模型自动训练单元 130 仅获得了一个线上机器学习模型，但是在线上机器学习模型上线并提供在线预测服务后，模型自动训练单元 130 会持续地得到更新的机器学习模型以供服务提供单元 140 选择用于提供在线预测服务的线上机器学习模型。

如图 3 中所示，每当完成图 3 中示出的一个操作，图 3 中央处的圆环图形中的相应图形部分就会相应地发生变化，以提示用户该过程已经完成，例如，当历史数据的真实结果上传成功后，图 3 中央处的圆环图形中的“反馈数据”图形部分就会相应地发生变化，以提示用户真实结果上传成功，即，真实结果收集单元 120 成功收集了历史数据的真实结果。

至此，已经完成了模型训练方案的探索、初始机器学习模型的训练以及机器学习模型的上线，下面将描述如何基于线上机器学习模型进行在线预测服务、数据回流、机器学习模型的更新等过程。

在步骤 S203 服务提供单元 140 将初始机器学习模型用作线上机器学习模型以提供在线预测服务（即，预估服务启动）后，用户可以利用预估服务的请求 API 地址进行预测服务请求，因此在步骤 S204，服务提供单元 140 可判断是否接收到包括预测数据的预测服务请求。

如果没有接收到预测服务请求，则继续在步骤 S204 进行判断。

如果接收到来自用户的预测服务请求，则进行到步骤 S205。例如，当用户想要预测其期望向消费者推荐的多种资讯是否会被消费者接受时，用户可向服务提供单元 140 提供包括预测数据的预测服务请求以获得用于确定向消费者推荐哪些资讯的依据，其中，预测数据可包括用户期望向消费者推荐的资讯以及消费者的基本属性信息（例如，资讯主题、资讯展示位置、消费者标识符、性别、年龄、身高、体重、兴趣爱好等），在这种情况下，可进行到步骤 S205。

在步骤 S205，服务提供单元 140 可自动保存包括在预测服务请求中的预测数据，并且数据收集单元 110 可持续从服务提供单元收集预测数据，其中，收集到的预测数据（在具有相应的真实结果之后）将由模型自动训练单元 130 用于得到更新的机器学习模型，随后将对此进行详细描述。通过步骤 S205，可实现数据的自动回流，从而为自动机器学习过程的不断循环提供了必要的数据来源。

在步骤 S206，服务提供单元 140 可响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

具体地讲，为了使线上机器学习模型能够自动地对预测数据进行特征工程，由模型自动训练单元 130 得到的机器学习模型所对应的模型文件的元数据中包括了相应的特征抽取过程信息，其中，特征抽取过程信息记载了如何对预测数据进行特征抽取，从而使得在步骤 S206 中，服务提供单元 140 可利用线上机器

学习模型所对应的文件中的特征抽取过程信息自动地对预测服务请求中的预测数据进行特征抽取，从而得到预测样本，并利用线上机器学习模型针对预测样本提供预测结果，即将该预测结果提供给发出预测服务请求的用户。

具体地讲，当包括在预测服务请求中的预测数据是包括资讯以及消费者的基本属性信息（例如，资讯主题、资讯展示位置、消费者标识符、性别、年龄、身高、体重、兴趣爱好等）时，服务提供单元 140 可利用线上机器学习模型得到这样的预测结果：预测会点击阅读的结果表示为 1，预测不会点击阅读的结果表示为 0，在此基础上，可以进一步提供对于某条资讯而言某个消费者可能点击该资讯的概率，例如，0.9、0.85、0.76 等。此后，服务提供单元 140 将该预测结果提供给发出预测服务请求的用户，而用户可根据接收的预测结果将消费者最可能点击的若干条资讯发送给消费者。

在步骤 S207，真实结果收集单元 120 可持续收集预测数据的真实结果，即，可持续地从用户收集真实结果。作为示例，用户可获取消费者是否实际点击了某条资讯（例如，某条被推荐资讯）的真实反馈，即，如果某条资讯被消费者点击浏览，则与包括该条资讯的预测数据相应的真实结果可被赋值为 1，如果某条资讯未被消费者点击浏览，则与包括该资讯的预测数据相应的真实结果可被赋值为 0。通过该过程可以实现预测数据的真实结果的自动回流，从而为自动机器学习过程的不断循环提供了必要的数据来源。

在步骤 S208，模型自动训练单元 130 可按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型，其中，配置的模型更新方案可以由模型自动训练单元 130 在训练出初始机器学习模型的模型训练方案基础上生成的，该模型更新方案还可包括以下项之中的至少一项：用于从预测数据中选取用于生成更新训练样本的预测数据的数据选取规则、用于由模型自动训练单元更新机器学习模型的模型更新周期、模型存储位置、更新资源自动配置方式，其中，模型更新方案中包括的以上项目都可以在系统中进行人为设置。

例如，可将模型更新周期设置为 1 星期，将数据选取规则设置为按时间范围选取数据（例如将数据范围设置为“前 7 天”），将模型存储位置设置为系统 100 内置的模型中心，并将更新资源自动配置方式设置为按照数据量结合规则来配置资源。

在此基础上，当根据模型更新周期需要进行模型更新时，在根据更新资源自动配置方式进行了系统资源配置之后，模型自动训练单元 130 可按照设置好的数据选择规则从收集的预测数据及其对应的真实结果中选择特定范围内的数据，对选取的数据进行特征抽取，优选地还可以进行特征组合，此后按照全量学习方式，根据模型更新方案中的模型训练方案，利用特征抽取后形成的更新训练样本进行模型训练以获得新的机器学习模型，最后将得到的新的机器学习模型存储在由模型存储位置所指示的地方。

可选地，当根据模型更新周期需要进行模型更新时，在根据更新资源自动配置方式进行了系统资源配置之后，模型自动训练单元 130 可根据配置的模型更新方案选取预测数据并进行特征抽取，优选地还可以进行特征组合，然后按照增量学习方式，利用特征抽取后形成的更新训练样本，对原来的机器学习模型进行增量学习，从而得到更新的机器学习模型，并将得到的更新的机器学习模型存储在由模型存储位置所指示的地方。

应注意，系统可被设计为仅按照全量学习方式产生更新的机器学习模型，或者被设计为仅按照增量学习方式产生更新的机器学习模型，或者被设计为可根据当前模型的预测效果来确定到底使用全量学习方式和增量学习方式中的哪一种来产生更新的机器学习模型，但是本公开并不限于此，任何可产生更新的机器学习模型的学习方式都可以被用于本公开。

通过步骤 S208，可得到更新的机器学习模型以供服务提供单元 140 进行选择来提供在线预测服务。

在步骤 S208 之后，进行到步骤 S209，服务提供单元 140 可按照配置的模型应用方案从由模型自动训练单元 130 获得的机器学习模型之中选择用于提供在线预测服务的线上机器学习模型。

具体地讲，服务提供单元 140 可根据模型应用方案中包括的模型选取规则从模型自动训练单元 130 所获得并存储的机器学习模型中选择一个或多个机器学习模型来作为线上机器学习模型，其中，模型选取规则可包括选择具有最高 AUC 的机器学习模型的规则、选择最新产生的机器学习模型的规则等规则。例如，服务提供单元 140 可根据 AUC 值从存储的机器学习模型中选择具有最高 AUC 的机器学习模型来作为线上机器学习模型。

在步骤 S209 之后，返回到步骤 S204，服务提供单元 140 继续判断是否接收到预测服务请求。

如果接收到预测服务请求,则服务提供单元 140 利用新选择的机器学习模型作为线上机器学习模型来提供在线预测服务。否则,服务提供单元 140 返回步骤 S204 继续判断。通过以上描述可知,一旦服务提供单元 140 利用所选择的机器学习模型提供在线预测服务,图 2 中的方法就可以一直按照参照图 2 所描述的循环过程形成一个闭环,从而可使得该闭环自动地持续运转和更新下去。

5 <实施例二>

<方法实施例>

图 9 示出根据本公开另一实施例的用于执行机器学习过程的方法的流程示意图。该方法可由至少一个计算装置执行,该至少一个计算装置可全部构建为本地装置,也可全部构建为云端装置(例如,云端服务器),还可同时包括本地装置和云端装置两者(例如,本地客户端与云端服务器两者)。

10 步骤 S9100,提供相互独立的第一操作入口和第二操作入口。

第一操作入口用于收集作为模型预测基础的行为数据,第二操作入口用于收集作为行为数据的真实结果的反馈数据。

15 行为数据涉及训练数据的特征部分,可以由用户按照不同路径来导入,比如将本地存储的数据上传系统,通过 HDFS 定时导入数据、通过 Kafka 实时流入数据等。这里,最初导入的数据会限定整个数据的 schema,后续再导入新的数据时,会进行 schema 核验,从而只接受同样 schema 内容的

20 数据表单。当导入的行为数据落盘时,会转换成相应数据组的特定格式,作为数据组中的各个数据切片,反馈数据也是同样的机制。
示例性地,如图 3 所示的图形用户界面,第一操作入口可以是对应的“进入”按钮 401,第二操作入口可以是对应的“进入”按钮 402,并且,第一操作入口和第二操作入口为两个相互独立的

25 操作入口,其中,第一操作入口用于收集行为数据,第二操作入口用于收集作为行为数据的真实结果的反馈数据。
步骤 S9200,获取并保存通过第一操作入口收集的行为数据和通过第二操作入口收集的反馈数据。

在本实施例中,该步骤 S9200 中获取并保存通过第一操作入口收集的行为数据和通过第二操作入口收集的反馈数据可以进一步包括如下步骤 S9210~S9230:

步骤 S9210,响应于针对第一操作入口和第二操作入口之中的任一操作入口的触发操作,提供至少一个数据导入路径供选择。

30 示例性地,如图 3 中所示的图形用户界面,为了导入行为数据,用户可点击该图形用户界面的左上方的与行为数据对应的“进入”按钮 401,然后进入图 4 所示的图形用户界面。在图 4 中,可通过三种导入方式来导入行为数据,即,将本地存储的数据上传系统,通过 HDFS 定时导入数据以及通过 Kafka 实时流入数据。

同样,为了导入反馈数据,用户可点击图 3 中所示的图形用户界面的右上方的与反馈数据对应的“进入”按钮 402,然后进入图 6 所示的图形用户界面。在图 6 中,同样可以通过三种导入方式来导入反馈数据,即,将本地存储的数据上传系统,通过 HDFS 定时导入数据以及通过 Kafka 实时流入数据。

35 步骤 S9220,从选择的数据导入路径导入行为数据或反馈数据。

可以理解的是,尽管对于行为数据和反馈数据,对应的页面中均示出了上述三种导入方式,但可根据具体应用情景而只启用其中的一种或多种,本实施例在此不做限定。

40 在本实施例中,该步骤 S9220 中从选择的数据导入路径导入行为数据或反馈数据可以进一步包括如下步骤 S9221~S9222:

步骤 S9221,在选择数据导入路径之后,提供对所导入的数据进行信息配置的配置界面。

45 示例性地,以导入行为数据为例,当用户选择了图 4 中所示的“本地上传数据”按钮之后,用户可从本地存储的数据中选择想要导入系统的行为数据。图 5 示出了已经被用户选中的将被导入系统的数据,而且,在图 5 中,提供对被用户选中的将被导入系统的数据进行信息配置的配置界面,具体的,该配置界面中可以提供目标数据、方案名称、文件首行为字段名、主键字段标记以及数据预览等配置信息,另外的,该配置界面中还提供有数据表的行数和列数,图 5 中数据表共有 37000 行和 37000 列。

步骤 S9222,根据通过配置界面输入的配置信息,导入行为数据或反馈数据。

如图 5 所示, 以导入行为数据为例, 为了使行为数据与反馈数据之间能够一一对应, 需要对行为数据的唯一 ID、数据列的类型等信息进行配置。当用户点击图 5 中的“下一步”之后可将该行为数据导入到系统中。

步骤 S9230, 保存导入的行为数据或反馈数据。

5 在本实施例中, 针对首次导入的行为数据或反馈数据, 以及, 针对后续导入的行为数据或反馈数据这两种不同情况, 提供两种不同的保存方式, 在此, 该步骤 S9230 中保存导入的行为数据或反馈数据可以进一步包括:

情况 1: 针对首次导入的行为数据或反馈数据执行结构提取, 并将行为数据或反馈数据保存为行为数据组或反馈数据组下的首个数据切片。

10 情况 2: 针对后续导入的行为数据或反馈数据执行结构核验, 并将核验通过的行为数据或反馈数据保存为行为数据组或反馈数据组下的后续数据切片。

步骤 S9300, 基于保存的行为数据和反馈数据, 利用至少一种模型算法来训练机器学习模型。

15 在本实施例中, 还提供分别与第一操作入口和第二操作入口相独立的第三操作入口, 第三操作入口用于进行关于模型训练的配置。具体的, 如图 3 所示的图形用户界面, 用户可点击该图形用户界面的右下方的与模型训练对应的“进入”按钮, 然后进入图 10 所示的图形用户界面。

在本实施例中, 该步骤 S9300 中基于保存的行为数据和反馈数据, 利用至少一种模型算法来训练机器学习模型可以进一步包括如下步骤 S9310~S9320:

步骤 S9310, 获取通过第三操作入口输入的配置信息。

20 通过第三操作入口输入的配置信息涉及探索模型训练方案的配置和/或在已有模型训练方案的基础上进行自学习的配置。

探索模型训练方案的配置包括行为数据选取规则、反馈数据选取规则、方案探索停止策略、数据自动拆分(训练/验证)、训练集所占比例以及随机种子中任意一项或者多项的信息的配置。

25 示例性地, 在图 10 中, 左侧展示的方案探索引擎用于进行探索模型训练方案的配置, 在用户首次进行方案探索的情况下, 可以是点击该图形用户界面的与进行探索模型训练方案的配置对应的“开启新探索”按钮以进行探索模型训练方案的配置。具体的, 可以是在点击“开启新探索”按钮之后, 然后进入图 11 所示的图形用户界面, 在图 11 中, 提供请选择行为数据切片范围, 用户可以是选择“该数据组的全部切片”, 也可以是“按数量范围调取切片”, 例如选择第 20 切片数据~第 200 切片数据, 还提供请选择反馈数据切片范围, 用户可以是选择“该数据组的全部切片”, 也可以是选择“按数量范围调取切片”, 例如选择第 20 切片数据~第 200 切片数据, 点击“下一步”之后, 然后进入图 12 所示的图形用户界面, 在图 30 12 中, 提供方案探索停止策略的配置, 用户可以是选择“手动停止”、“达到 AUC”、“达到训练时长”以及“达到训练轮数”, 还提供数据自动拆分(训练/验证)的配置, 用户可以是选择“按比例拆分”、“按规则拆分”以及“先排序后拆分数据”, 还提供训练集所占比例, 用户可以是设置比例为“0.8”等。

35 在已有模型训练方案的基础上进行自学习的配置包括手动进行一次自学习以及配置定时自学习计划的配置中的至少一项, 其中, 手动进行一次自学习的配置包括数据来源和数据切片选取的信息的配置; 定时自学习计划的配置包括自学习周期、自学习数据以及自学习结果中任意一项或者多项的信息的配置。

40 示例性地, 在图 10 中, 右侧展示的模型工厂用于在已有模型训练方案的基础上进行自学习的配置, 可以是点击该图形用户界面的与在已有模型训练方案的基础上进行自学习对应的“选择方案”按钮, 以进行在已有模型训练方案的基础上进行自学习的配置。具体的, 可以是在点击“选择方案”按钮之后, 然后进入图 13 所示的图形用户界面, 在图 13 中, 提供“手动进行一次自学习”的配置按钮, 以及提供“配置定时自学习计划”的配置按钮, 用户可以是点击“手动进行一次自学习”按钮之后, 然后进入图 14 所示的图形用户界面, 在图 14 中, 用户可以选择数据来源, 或者是数据切片选取; 也可以是在点击“配置定时自学习计划”之后, 然后进入图 15 所示的图形用户界面, 在图 15 中, 提供自学习周期的配置, 用户可以是选择运行模式为“单次运行”、“循环运行”以及“crontab 表达式”, 以及选择任务开始时间为“2019-06-17 45 11:38:43”, 还提供自学习数据配置, 用户可以是进行数据来源、数据切片、模型命名结果以及任务超时时长的选择等。

步骤 S9320, 根据通过第三操作入口输入的配置信息, 将保存的行为数据和反馈数据拼接成训练数据, 通过对训练数据进行特征工程(例如, 特征抽取)来生成训练样本, 利用至少一种模型算

法，基于训练样本来训练机器学习模型。

在本实施例中，基于保存的行为数据和反馈数据生成训练样本并利用训练样本训练出机器学习模型。可以是将保存的行为数据的 ID 和反馈数据的 ID 进行拼接以形成训练数据，并对训练数据进行特征工程（例如，包括特征抽取和/或特征组合等）来生成训练样本。具体地讲，拼接后得到的训练数据可以以数据表的形式进行存储。通常，数据表的一行可对应于一条数据记录，数据表的一列可对应于一个属性字段。数据表中的每条数据记录可包括一个或更多个属性信息（即，属性字段）。属性字段可用于形成特征，而特征是构成训练样本的组成部分，这里，特征可以是属性字段本身或属性字段的局部、属性字段的组合、或属性字段经过处理（或运算）而得到的结果。不同特征之间还可以进一步结合。

5 示例性地，图 7 中间部分示出的有向无环图（DAG 图）示出了 6 个节点：“反馈数据”节点、“行为数据”节点、“数据拆分”节点、“特征工程”节点、“LR（逻辑回归）算法”节点、“GBDT（梯度提升决策树）算法”节点、“HE-TreeNet（高维离散嵌入式的树网络）算法”节点和“NN（神经网络）算法”节点。应注意，图 7 示出了 4 种具体预置算法，但是这仅是示例性说明，本公开并不对预置算法的数量和具体算法进行限制。

10 参见图 7，经由 DAG 图中的“数据拆分”节点处的相应配置，可将行为数据和反馈数据拼接后的训练数据拆分为训练集和验证集。此后，经由 DAG 图中的“特征工程”节点处的相应配置，可对训练集和验证集分别进行自动特征生成来抽取至少一个特征以生成训练样本。对应于 DAG 图中的最下层的三个节点（即“LR 算法”节点、“GBDT 算法”节点、“HE-TreeNet 算法”节点和“NN 算法”节点）处，利用训练样本对这四个预置算法分别进行至少一轮训练，进而训练出相应的多个机器学习模型。

20 进一步地，在本实施例中，可以是将以下过程可视化地展示给用户：根据涉及探索模型训练方案的配置的配置信息，将保存的行为数据和反馈数据拼接成训练数据，通过对训练数据进行特征工程（例如，特征抽取）来生成训练样本，利用至少一种模型算法，基于训练样本来训练机器学习模型。

25 该将以下过程可视化地展示给用户的步骤包括展示以下项目中的至少一项：向用户展示将保存的行为数据和反馈数据拼接成训练数据的数据处理进度、向用户展示通过对训练数据进行特征抽取来生成训练样本的过程中涉及的特征维度和/或特征重要性、向用户展示模型探索实验的轮数、已运行时间和/或效果指标、向用户展示模型训练的算法及其效果指标、向用户展示探索模型训练方案的流程示意图。

30 根据本实施例的方法，其分别提供收集行为数据的操作入口，以及，收集反馈数据的操作入口，以将行为数据和反馈数据分别导入系统，方便用户通过容易理解的交互方式来完成机器学习模型的自动训练过程。

在一个实施例中，还提供分别与第一操作入口和第二操作入口相独立的第四操作入口，第四操作入口用于进行关于利用机器学习模型提供预测服务的配置。

35 示例性地，如图 3 所示的图形用户界面，用户可点击该图形用户界面的左下方的与模型应用对应的“进入”按钮，然后进入图 16 所示的图形用户界面。

在本实施例中，本公开用于执行机器学习过程的方法还包括以下步骤 S9400~S9500：

步骤 S9400，获取通过第四操作入口输入的配置信息。

40 通过第四操作入口输入的配置信息涉及利用机器学习模型提供在线预测服务和/或批量预测服务。在图 16 中，分别提供与在线预测服务对应的“在线预估”按钮，以及与批量预测服务对应的“批量预估”按钮。

涉及在线预测服务的配置信息包括更改服务的配置，例如，选择需要上线的模型和分配资源的信息的配置中的至少一项；以及，涉及批量预测服务的配置信息包括编辑预估服务的配置，例如，选择需要上线的机器学习模型的信息的配置。

步骤 S9500，基于通过第四操作入口输入的配置信息，利用机器学习模型提供预测服务。

45 在本实施例中，通过第四操作入口输入的配置信息涉及利用机器学习模型提供在线预测服务和/或批量预测服务，在此，该步骤 S9500 基于通过第四操作入口输入的配置信息，利用机器学习模型提供预测服务可以进一步包括：

基于通过第四操作入口输入的涉及在线预测服务的配置信息和/或涉及批量预测服务的配置信

息，利用机器学习模型提供在线预测服务和/或批量预测服务。

继续上述示例，可以根据模型效果等因素从以上训练出的多个机器学习模型中选取一个或多个机器学习模型作为提供预测服务的机器学习模型，其中，如果多个机器学习模型被选取，则可综合这些机器学习模型的预测结果来得到向外提供的预测结果。

- 5 在本实施例中，基于通过第四操作入口输入的涉及在线预测服务的配置信息，利用机器学习模型提供在线预测服务，包括：

步骤 S9510，通过配置信息中设定的 API 地址接收包括预测数据的预测服务请求。

用户可以利用预估服务的请求 API 地址进行预测服务请求。

- 10 步骤 S9520，响应于接收到的预测服务请求，利用机器学习模型得到针对预测数据的预测结果，并将预测结果通过 API 地址发出。

在本实施例中，通过第四操作入口输入的涉及在线预测服务的配置信息还包括预测数据的自动回流的开关状态，在此，本公开用于执行机器学习过程的方法还包括：

在开关状态为开启的情况下，将包括在预测服务请求中的预测数据保存在相应的行为数据组。

- 15 在图 17 所示的图形用户界面中，提供是否保存您的预估请求的选择选项，在用户选择“关闭自动回流”按钮的情况下，则不会将在预测服务请求中的预测数据保存在相应的行为数据组，在用户选择“开启自动回流”按钮的情况下，则会将包括在预测服务请求中的预测数据保存在相应的行为数据组。通过该过程可以实现行为数据的自动回流，从而为自动机器学习过程的不断循环提供了必要的数据来源。

在一个实施例中，所有操作入口被提供在同一交互界面上。

- 20 作为示例，第一操作入口、第二操作入口、第三操作入口和第四操作入口均被提供在如图 3 所示的图形界面中，其中，第一操作入口可以是该图形用户界面的左上方的与行为数据对应的“进入”按钮 401，也可以是该图形用户界面的中央处的圆环图形中的“行为数据”图形；第二操作入口可以是该图形用户界面的右上方的与反馈数据对应的“进入”按钮 402，也可以是该图形用户界面的中央处的圆环图形中的“反馈数据”图形；第三操作入口可以是该图形用户界面的右下方的与模型训练对应的“进入”按钮，也可以是该图形用户界面的中央处的圆环图形中的“模型训练”图形，以及，第四操作入口可以是该图形用户界面的左下方的与模型应用对应的“进入”按钮，也可以是该图形用户界面的中央处的圆环图形中的“模型应用”图形。

在一个实施例中，本公开用于执行机器学习过程的方法还包括以下步骤 S10011~S10013：

步骤 S10011，在交互界面上，提供对应每一操作入口的信息展示区。

- 30 在图 3 所示的图形界面中，第一操作入口可以是该图形用户界面的左上方的与行为数据对应的“进入”按钮 401，对应第一操作入口的信息展示区可以是“进入”按钮 401 上方显示的信息；第二操作入口可以是该图形用户界面的右上方的与反馈数据对应的“进入”按钮 402，对应第二操作入口的信息展示区可以是“进入”按钮 402 上方显示的信息；第三操作入口可以是该图形用户界面的右下方的与模型训练对应的“进入”按钮，对应第三操作入口的信息展示区可以是与模型训练对应的“进入”按钮上方显示的信息，以及，第四操作入口可以是该图形用户界面的左下方的与模型应用对应的“进入”按钮，对应第四操作入口的信息展示区可以是与模型应用对应的“进入”按钮上方显示的信息。

步骤 S10012，获取对应每一操作入口的当前操作状态信息。

步骤 S10013，配置对应每一操作入口的信息展示区，展示对应操作入口的当前操作状态信息。

- 40 如图 3 中所示，例如，当未上传行为数据时，该图形用户界面的左上方的与行为数据对应的“进入”按钮 401 的右上方至少显示有“数据待接入”；又例如，当未上传反馈数据时，该图形用户界面的左上方的与行为数据对应的“进入”按钮 401 的右上方至少显示有“数据未收集”。

此外，所述当前操作状态信息还可包括各个操作所涉及的操作对象（例如，行为数据、反馈数据、模型方案、预估请求）、操作内容和/或操作结果的信息。

在一个实施例中，本公开用于执行机器学习过程的方法还包括以下步骤 S10021~S10023：

- 45 步骤 S10021，对于每一操作入口，分别提供与各自相对应的进度指示条。

在本实施例中，可以是设置每一操作入口同时作为与各自相对应的进度指示条使用。示例性地，在图 3 所示的图形用户界面中，第一操作入口可以是该图形用户界面的中央处的圆环图形中的“行为数据”图形，该“行为数据”图形可直接作为第一操作入口对应的进度指示条；第二操作入口可

以是该图形用户界面的中央处的圆环图形中的“反馈数据”图形，该“反馈数据”图形可直接作为第二操作入口对应的进度指示条；第三操作入口可以是该图形用户界面的中央处的圆环图形中的“模型训练”图形，该“模型训练”图形可直接作为第三操作入口对应的进度指示条，以及，第四操作入口可以是该图形用户界面的中央处的圆环图形中的“模型应用”图形，该“模型应用”图形可直接作为第四操作入口对应的进度指示条。

步骤 S10022，对于每一操作入口，检测执行对应操作的当前进度。

步骤 S10023，根据检测到的当前进度，控制对应进度指示条的显示状态。

如图 3 中所示，每当完成图 3 中示出的一个操作，图 3 中央处的圆环图形中的相应图形部分就会相应地发生变化，以提示用户该过程已经完成，例如，当行为数据上传成功后，图 3 中央处的圆环图形中的“行为数据”图形部分就会相应地发生变化，以提示用户行为数据上传成功；又例如，当反馈数据上传成功后，图 3 中央处的圆环图形中的“反馈数据”图形部分就会相应地发生变化，以提示用户反馈数据上传成功。

<系统实施例>

在本实施例中，还提供一种用于执行机器学习过程的系统 9000，如图 18 所示，用于执行机器学习过程的系统 9000 包括交互单元 9100、数据收集单元 9200、真实结果收集单元 9300 以及模型自动训练单元 9400。

交互单元 9100，用于提供相互独立的第一操作入口和第二操作入口，其中，所述第一操作入口用于收集作为模型预测基础的行为数据，所述第二操作入口用于收集作为行为数据的真实结果的反馈数据。

数据收集单元 9200，用于获取并保存通过所述第一操作入口收集的行为数据。

真实结果收集单元 9300，用于获取并保存通过所述第二操作入口收集的反馈数据。

模型自动训练单元 9400，用于基于保存的所述行为数据和反馈数据，利用至少一种模型算法来训练机器学习模型。

在一个实施例中，数据收集单元 9200，还用于响应于针对所述第一操作入口的触发操作，提供至少一个数据导入路径供选择；从选择的数据导入路径导入行为数据；以及保存导入的所述行为数据。

真实结果收集单元 9300，还用于响应于针对所述第二操作入口的触发操作，提供至少一个数据导入路径供选择；从选择的数据导入路径导入反馈数据；以及保存导入的所述反馈数据。

在一个实施例中，数据收集单元 9200，还用于在选择数据导入路径之后，提供对所导入的数据进行信息配置的配置界面；以及根据通过所述配置界面输入的配置信息，导入行为数据。

真实结果收集单元 9300，还用于在选择数据导入路径之后，提供对所导入的数据进行信息配置的配置界面；以及，根据通过所述配置界面输入的配置信息，导入反馈数据。

在一个实施例中，数据收集单元 9200，还用于针对首次导入的所述行为数据执行结构提取，并将所述行为数据保存为行为数据组下的首个数据切片；以及，针对后续导入的所述行为数据执行结构核验，并将核验通过的行为数据保存为行为数据组下的后续数据切片。

真实结果收集单元 9300，还用于针对首次导入的所述反馈数据执行结构提取，并将所述反馈数据保存为反馈数据组下的首个数据切片；以及，针对后续导入的所述反馈数据执行结构核验，并将核验通过的反馈数据保存为反馈数据组下的后续数据切片。

在一个实施例中，交互单元 9100，还用于提供分别与所述第一操作入口和所述第二操作入口相独立的第三操作入口，所述第三操作入口用于进行关于模型训练的配置。

模型自动训练单元 9400，还用于获取通过所述第三操作入口输入的配置信息；根据通过所述第三操作入口输入的配置信息，将保存的行为数据和反馈数据拼接成训练数据，通过对训练数据进行特征抽取来生成训练样本，利用至少一种模型算法，基于训练样本来训练机器学习模型。

在一个实施例中，所述通过所述第三操作入口输入的配置信息涉及探索模型训练方案的配置和/或在已有模型训练方案的基础上进行自学习的配置。

在一个实施例中，所述模型自动训练单元 9400，还用于将以下过程可视化地展示给用户：根据涉及探索模型训练方案的配置的配置信息，将保存的行为数据和反馈数据拼接成训练数据，通过对训练数据进行特征抽取来生成训练样本，利用至少一种模型算法，基于训练样本来训练机器学习

模型。

在一个实施例中，交互单元 9100，还用于提供分别与所述第一操作入口和所述第二操作入口相独立的第四操作入口，所述第四操作入口用于进行关于利用机器学习模型提供预测服务的配置。

如图 19 所示，该用于执行机器学习过程的系统 9000 还可以包括服务提供单元 9500。

5 该服务提供单元 9500，用于基于通过所述第四操作入口输入的配置信息，利用机器学习模型提供预测服务。

在一个实施例中，所述通过所述第四操作入口输入的配置信息涉及利用机器学习模型提供在线预测服务和/或批量预测服务。

10 在一个实施例中，该服务提供单元 9500，还用于基于通过所述第四操作入口输入的涉及在线预测服务的配置信息和/或涉及批量预测服务的配置信息，利用机器学习模型提供在线预测服务和/或批量预测服务。

在一个实施例中，该服务提供单元 9500，还用于通过所述配置信息中设定的 API 地址接收包括预测数据的预测服务请求；响应于接收到的所述预测服务请求，利用机器学习模型得到针对预测数据的预测结果，并将预测结果通过所述 API 地址发出。

15 在一个实施例中，所述通过所述第四操作入口输入的涉及在线预测服务的配置信息还包括预测数据的自动回流的开关状态。

该服务提供单元 9500，还用于在所述开关状态为开启的情况下，将包括在所述预测服务请求中的预测数据保存在相应的行为数据组。

在一个实施例中，所有操作入口被提供在同一交互界面上。

20 在一个实施例中，该用于执行机器学习过程的系统 9000 还可以包括操作状态展示单元。

该操作状态展示单元 9600，用于在所述交互界面上，提供对应每一操作入口的信息展示区；获取对应每一操作入口的当前操作状态信息；配置所述对应每一操作入口的信息展示区，展示对应操作入口的当前操作状态信息。

在一个实施例中，该用于执行机器学习过程的系统 9000 还可以包括进度展示单元。

25 该进度展示单元 9700，用于对于每一操作入口，分别提供与各自相对应的进度指示条；对于每一操作入口，检测执行对应操作的当前进度；根据检测到的所述当前进度，控制对应进度指示条的显示状态。

在一个实施例中，该进度展示单元 9700，还用于设置每一操作入口同时作为与各自相对应的进度指示条使用。

30 <装置实施例>

在本实施例中，还提供一种用于执行机器学习过程的计算装置 10000。

如图 20 所示，用于执行机器学习过程的计算装置 10000 可以包括用于执行机器学习过程的系统，例如可以是图 1 示出的用于执行机器学习过程的系统 100，也可以是图 18 或图 19 示出的用于执行机器学习过程的系统 9000，在此不做限定。

35 在另一个实施例中，如图 21 所示，用于执行机器学习过程的计算装置 10000 还可以包括处理器 10100 和存储部件 10200，存储部件 10200 中存储有计算机可执行指令集合，当所述计算机可执行指令被所述处理器 10100 执行时，促使所述处理器 10100 执行根据本公开第二实施例的用于执行机器学习过程的方法。

<计算机可读存储介质>

40 在本实施例中，还提供一种计算机可读存储介质，其上存储有计算机程序，计算机程序在被处理器执行时实现如本公开任意实施例的阅读信息的处理方法。

根据本公开示例性实施例的用于执行机器学习过程的系统所包括的单元可被分别配置为执行特定功能的软件、硬件、固件或上述项的任意组合。例如，这些单元可对应于专用的集成电路，也可对应于纯粹的软件代码，还可对应于软件与硬件相结合的模块。此外，这些单元所实现的一个或多个功能也可由物理实体设备（例如，处理器、客户端或服务器等）中的组件来统一执行。

45 应理解，根据本公开示例性实施例的用于执行机器学习过程的方法可通过记录在计算可读介质上的程序来实现，例如，根据本公开的示例性实施例，可提供一种存储指令的计算机可读存储介质，其中，当所述指令被至少一个计算装置运行时，促使所述至少一个计算装置执行所述用于执行机器

学习过程的方法中的各个步骤的计算机程序。

上述计算机可读介质中的计算机程序可在诸如客户端、主机、代理装置、服务器等计算机设备中部署的环境中运行，应注意，所述计算机程序还可用于执行在执行上述步骤时执行更为具体的处理，这些进一步处理的内容已经参照图 2 至图 8 进行了描述，这里为了避免重复将不再进行赘述。

5 应注意，根据本公开示例性实施例的用于执行机器学习过程的系统可完全依赖计算机程序的运行来实现相应的功能，即，各个装置与计算机程序的功能架构中与各步骤相应，使得整个系统通过专门的软件包（例如，lib 库）而被调用，以实现相应的功能。

另一方面，根据本公开示例性实施例的用于执行机器学习过程的系统所包括的各个单元也可以通过硬件、软件、固件、中间件、微代码或其任意组合来实现。当以软件、固件、中间件或微代码实现时，用于执行相应操作的程序代码或者代码段可以存储在诸如存储介质的计算机可读介质中，使得处理器可通过读取并运行相应的程序代码或者代码段来执行相应的操作。

10 例如，本公开的示例性实施例还可以实现为计算装置，该计算装置包括存储部件和处理器，存储部件中存储有计算机可执行指令集合，当所述计算机可执行指令集合被所述处理器执行时，执行用于执行机器学习过程的方法。例如，一种包括至少一个计算装置和至少一个存储指令的存储装置的系统，其中，所述指令在被所述至少一个计算装置运行时，促使所述至少一个计算装置执行以上描述的用于执行机器学习过程的方法。

具体说来，所述计算装置可以部署在服务器或客户端中，也可以部署在分布式网络环境中的节点装置上。此外，所述计算装置可以是 PC 计算机、平板装置、个人数字助理、智能手机、web 应用或其他能够执行上述指令集合的装置。

20 这里，所述计算装置并非必须是单个的计算装置，还可以是任何能够单独或联合执行上述指令（或指令集）的装置或电路的集合体。计算装置还可以是集成控制系统或系统管理器的一部分，或者可被配置为与本地或远程（例如，经由无线传输）以接口互联的便携式电子装置。

在所述计算装置中，处理器可包括中央处理器（CPU）、图形处理器（GPU）、可编程逻辑装置、专用处理器系统、微控制器或微处理器。作为示例而非限制，处理器还可包括模拟处理器、数字处理器、微处理器、多核处理器、处理器阵列、网络处理器等。

根据本公开示例性实施例的用于执行机器学习过程的方法中所描述的某些操作可通过软件方式来实现，某些操作可通过硬件方式来实现，此外，还可通过软硬件结合的方式来实现这些操作。

30 处理器可运行存储在存储部件之一中的指令或代码，其中，所述存储部件还可以存储数据。指令和数据还可经由网络接口装置而通过网络被发送和接收，其中，所述网络接口装置可采用任何已知的传输协议。

存储部件可与处理器集成为一体，例如，将 RAM 或闪存布置在集成电路微处理器等之内。此外，存储部件可包括独立的装置，诸如，外部盘驱动、存储阵列或任何数据库系统可使用的其他存储装置。存储部件和处理器可在操作上进行耦合，或者可例如通过 I/O 端口、网络连接等互相通信，使得处理器能够读取存储在存储部件中的文件。

35 此外，所述计算装置还可包括视频显示器（诸如，液晶显示器）和用户交互接口（诸如，键盘、鼠标、触摸输入装置等）。计算装置的所有组件可经由总线和/或网络而彼此连接。

根据本公开示例性实施例的用于执行机器学习过程的方法所涉及的操作可被描述为各种互联或耦合的功能块或功能示图。然而，这些功能块或功能示图可被均等地集成为单个的逻辑装置或按照非确切的边界进行操作。

40 例如，如上所述，根据本公开示例性实施例的用于执行机器学习过程的计算装置可包括存储部件和处理器，其中，存储部件中存储有计算机可执行指令集合，当所述计算机可执行指令集合被所述处理器执行时，执行用于执行机器学习过程的方法中的各个步骤。

权 利 要 求 书

- 1、一种包括至少一个计算装置和至少一个存储指令的存储装置的系统，其中，所述指令在被所述至少一个计算装置运行时，促使所述至少一个计算装置执行用于执行机器学习过程的以下步骤，包括：
- 5 持续收集预测数据；
持续收集预测数据的真实结果；
按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及
- 10 按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。
- 2、如权利要求 1 所述的系统，其中，所述指令在被所述至少一个计算装置运行时，将促使所述至少一个计算装置还执行以下步骤：
- 15 自动保存包括在预测服务请求中的预测数据，
持续收集自动保存的预测数据。
- 3、如权利要求 2 所述的系统，其中，所述指令在被所述至少一个计算装置运行时，将促使所述至少一个计算装置还执行以下步骤：
- 20 收集历史数据；
收集历史数据的真实结果；
根据自动机器学习技术基于收集的历史数据及其对应的真实结果生成初始训练样本并利用初始训练样本训练出初始机器学习模型，并且，在初始机器学习模型的基础上，按照配置的模型更新方案来利用更新训练样本持续得到更新的机器学习模型。
- 4、如权利要求 3 所述的系统，其中，在训练出初始机器学习模型的模型训练方案基础上生成配置的模型更新方案。
- 25 5、如权利要求 3 所述的系统，其中，自动机器学习技术涉及以下项中的至少一项：用于将历史数据拆分为训练数据和验证数据的自动数据拆分、用于按照数据字段类型和数据分布之中的至少一个对训练数据和验证数据进行特征抽取的自动特征生成、用于按照特征重要性来确定组合特征的自动特征组合、用于按照预设调参方式进行参数调优的自动调参、用于根据预测效果来确定由一个或多个训练出的模型来提供预测结果的模型自动选择、以及根据预设规则对资源的自动配置。
- 30 6、如权利要求 1 所述的系统，其中，所述指令在被所述至少一个计算装置运行时，将促使所述至少一个计算装置还执行以下步骤：将得到的机器学习模型所对应的模型文件的元数据中包括相应的特征抽取过程信息。
- 7、如权利要求 6 所述的系统，其中，所述指令在被所述至少一个计算装置运行时，将促使所述至少一个计算装置还执行以下步骤：利用线上机器学习模型所对应的文件中的特征抽取过程信息自动地对预测服务请求中的预测数据进行特征抽取以得到预测样本，并利用线上机器学习模型针对预测样本提供预测结果。
- 35 8、如权利要求 1 所述的系统，其中，模型更新方案还包括以下项之中的至少一项：用于从预测数据中选取用于生成更新训练样本的预测数据的数据选取规则、用于由模型自动训练单元更新机器学习模型的模型更新周期、模型存储位置、更新资源自动配置方式。
- 40 9、如权利要求 3 所述的系统，其中，模型应用方案包括：用于从机器学习模型之中选择线上机器学习模型的模型选取规则和应用资源自动配置方式之中的至少一个。
- 10、一种存储指令的计算机可读存储介质，其中，当所述指令被至少一个计算装置运行时，促使所述至少一个计算装置执行：
- 45 持续收集预测数据；
持续收集预测数据的真实结果；
按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用

更新训练样本持续得到更新的机器学习模型；以及

按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

5 11、一种用于执行机器学习过程的系统，包括：

数据收集单元，用于持续收集预测数据；

真实结果收集单元，用于持续收集预测数据的真实结果；

模型自动训练单元，用于按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及

10 服务提供单元，用于按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

12、一种由至少一个计算装置执行的用于执行机器学习过程的方法，包括：

持续收集预测数据；

15 持续收集预测数据的真实结果；

按照配置的模型更新方案，基于收集的预测数据及其对应的真实结果生成更新训练样本并利用更新训练样本持续得到更新的机器学习模型；以及

按照配置的模型应用方案从机器学习模型之中选择用于提供在线预测服务的线上机器学习模型，并响应于包括预测数据的预测服务请求，利用线上机器学习模型针对包括在预测服务请求中的预测数据提供预测结果。

20 13、如权利要求 12 所述的方法，还包括：

自动保存包括在预测服务请求中的预测数据；

持续收集自动保存的预测数据。

14、如权利要求 13 所述的方法，还包括：

25 收集历史数据；

收集历史数据的真实结果；

根据自动机器学习技术基于收集的历史数据及其对应的真实结果生成初始训练样本并利用初始训练样本训练出初始机器学习模型，并且，在初始机器学习模型的基础上，按照配置的模型更新方案来利用更新训练样本持续得到更新的机器学习模型。

30 15、如权利要求 14 所述的方法，其中，在训练出初始机器学习模型的模型训练方案基础上生成配置的模型更新方案。

16、如权利要求 14 所述的方法，其中，自动机器学习技术涉及以下项中的至少一项：用于将历史数据拆分为训练数据和验证数据的自动数据拆分、用于按照数据字段类型和数据分布之中的至少一个对训练数据和验证数据进行特征抽取的自动特征生成、用于按照特征重要性来确定组合特征的自动特征组合、用于按照预设调参方式进行参数调优的自动调参、用于根据预测效果来确定由一个或多个训练出的模型来提供预测结果的模型自动选择、以及根据预设规则对资源的自动配置。

35 17、如权利要求 12 所述的方法，还包括：将得到的机器学习模型所对应的模型文件的元数据中包括相应的特征抽取过程信息。

18、如权利要求 17 所述的方法，还包括：利用线上机器学习模型所对应的文件中的特征抽取过程信息自动地对预测服务请求中的预测数据进行特征抽取以得到预测样本，并利用线上机器学习模型针对预测样本提供预测结果。

19、如权利要求 12 所述的方法，其中，模型更新方案还包括以下项之中的至少一项：用于从预测数据中选取用于生成更新训练样本的预测数据的数据选取规则、用于由模型自动训练单元更新机器学习模型的模型更新周期、模型存储位置、更新资源自动配置方式。

45 20、如权利要求 14 所述的方法，其中，模型应用方案包括：用于从机器学习模型之中选择线上机器学习模型的模型选取规则和应用资源自动配置方式之中的至少一个。

21、一种由至少一个计算装置执行的用于执行机器学习过程的方法，包括：

提供相互独立的第一操作入口和第二操作入口，其中，所述第一操作入口用于收集作为模型预

测基础的行为数据，所述第二操作入口用于收集作为行为数据的真实结果的反馈数据；

获取并保存通过所述第一操作入口收集的行为数据和通过所述第二操作入口收集的反馈数据；
基于保存的所述行为数据和反馈数据，利用至少一种模型算法来训练机器学习模型。

22、根据权利要求 21 所述的方法，其中，所述获取并保存通过所述第一操作入口收集的行为
5 数据和通过所述第二操作入口收集的反馈数据，包括：

响应于针对所述第一操作入口和所述第二操作入口之中的任一操作入口的触发操作，提供至少
一个数据导入路径供选择；

从选择的数据导入路径导入行为数据或反馈数据；以及
保存导入的所述行为数据或所述反馈数据。

23、根据权利要求 22 所述的方法，其中，所述从选择的数据导入路径导入行为数据或反馈数
10 据，包括：

在选择数据导入路径之后，提供对所导入的数据进行信息配置的配置界面；

根据通过所述配置界面输入的配置信息，导入行为数据或反馈数据。

24、根据权利要求 22 所述的方法，所述保存导入的所述行为数据或所述反馈数据，包括：

15 针对首次导入的所述行为数据或所述反馈数据执行结构提取，并将所述行为数据或所述反馈数
据保存为行为数据组或反馈数据组下的首个数据切片；以及，

针对后续导入的所述行为数据或所述反馈数据执行结构核验，并将核验通过的行为数据或反馈
数据保存为行为数据组或反馈数据组下的后续数据切片。

25、根据权利要求 21 所述的方法，所述方法还包括：

20 提供分别与所述第一操作入口和所述第二操作入口相独立的第三操作入口，所述第三操作入口
用于进行关于模型训练的配置；

所述基于保存的所述行为数据和反馈数据，利用至少一种模型算法来训练机器学习模型，包括：
获取通过所述第三操作入口输入的配置信息；

25 根据通过所述第三操作入口输入的配置信息，将保存的行为数据和反馈数据拼接成训练数据，
通过对训练数据进行特征抽取来生成训练样本，利用至少一种模型算法，基于训练样本来训练机器
学习模型。

26、根据权利要求 25 所述的方法，其中，所述通过所述第三操作入口输入的配置信息涉及探
索模型训练方案的配置和在已有模型训练方案的基础上进行自学习的配置之中的至少一个。

27、根据权利要求 26 所述的方法，所述方法还包括：将以下过程可视化地展示给用户：根据
30 涉及探索模型训练方案的配置的配置信息，将保存的行为数据和反馈数据拼接成训练数据，通过
对训练数据进行特征抽取来生成训练样本，利用至少一种模型算法，基于训练样本来训练机器学
习模型。

28、根据权利要求 21 所述的方法，所述方法还包括：

35 还提供分别与所述第一操作入口和所述第二操作入口相独立的第四操作入口，所述第四操作入
口用于进行关于利用机器学习模型提供预测服务的配置；

获取通过所述第四操作入口输入的配置信息；

基于通过所述第四操作入口输入的配置信息，利用机器学习模型提供预测服务。

29、根据权利要求 28 所述的方法，其中，所述通过所述第四操作入口输入的配置信息涉及利
40 用机器学习模型提供在线预测服务和批量预测服务之中的至少一个，并且，

所述基于通过所述第四操作入口输入的配置信息，利用机器学习模型提供预测服务，包括：

基于通过所述第四操作入口输入的涉及在线预测服务的配置信息和涉及批量预测服务的配置
信息之中的至少一个，利用机器学习模型提供在线预测服务和批量预测服务之中的至少一个。

30、根据权利要求 29 所述的方法，其中，所述基于通过所述第四操作入口输入的涉及在线预
测服务的配置信息，利用机器学习模型提供在线预测服务，包括：

45 通过所述配置信息中设定的 API 地址接收包括预测数据的预测服务请求；

响应于接收到的所述预测服务请求，利用机器学习模型得到针对预测数据的预测结果，并将预
测结果通过所述 API 地址发出。

31、根据权利要求 30 所述的方法，其中，所述通过所述第四操作入口输入的涉及在线预测服

务的配置信息还包括预测数据的自动回流的开关状态，所述方法还包括：

在所述开关状态为开启的情况下，将包括在所述预测服务请求中的预测数据保存在相应的行为数据组。

32、根据权利要求 21 所述的方法，其中，所有操作入口被提供在同一交互界面上。

5 33、根据权利要求 32 所述的方法，所述方法还包括：

在所述交互界面上，提供对应每一操作入口的信息展示区；

获取对应每一操作入口的当前操作状态信息；

配置所述对应每一操作入口的信息展示区，展示对应操作入口的当前操作状态信息。

34、根据权利要求 32 所述的方法，所述方法还包括：

10 对于每一操作入口，分别提供与各自相对应的进度指示条；

对于每一操作入口，检测执行对应操作的当前进度；

根据检测到的所述当前进度，控制对应进度指示条的显示状态。

35、根据权利要求 34 所述的方法，其中，所述对于每一操作入口，分别提供与各自相对应的进度指示条，包括：

15 设置每一操作入口同时作为与各自相对应的进度指示条使用。

36、一种用于执行机器学习过程的系统，包括：

交互单元，用于提供相互独立的第一操作入口和第二操作入口，其中，所述第一操作入口用于收集作为模型预测基础的行为数据，所述第二操作入口用于收集作为行为数据的真实结果的反馈数据；

20 数据收集单元，用于获取并保存通过所述第一操作入口收集的行为数据；

真实结果收集单元，用于获取并保存通过所述第二操作入口收集的反馈数据；以及，

模型自动训练单元，用于基于保存的所述行为数据和反馈数据，利用至少一种模型算法来训练机器学习模型。

25 37、一种包括至少一个计算装置和至少一个存储指令的存储装置的系统，其中，所述指令在被所述至少一个计算装置运行时，促使所述至少一个计算装置执行根据权利要求 21 至 35 中任一项所述的方法。

38、一种存储指令的计算机可读存储介质，其中，当所述指令被至少一个计算装置运行时，促使所述至少一个计算装置执行权利要求 21 至 35 中任一项所述的方法。

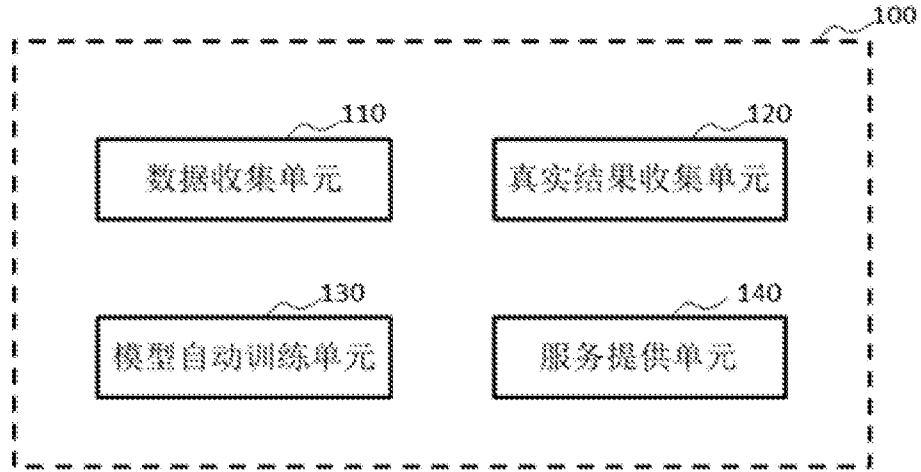


图 1

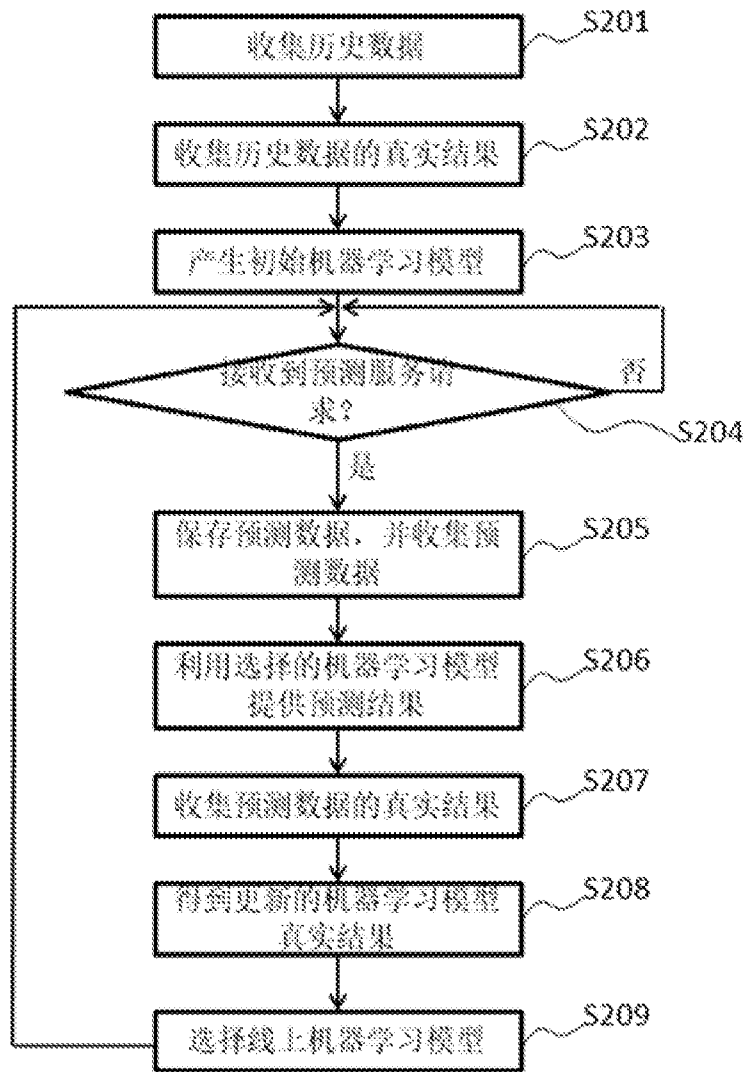


图 2

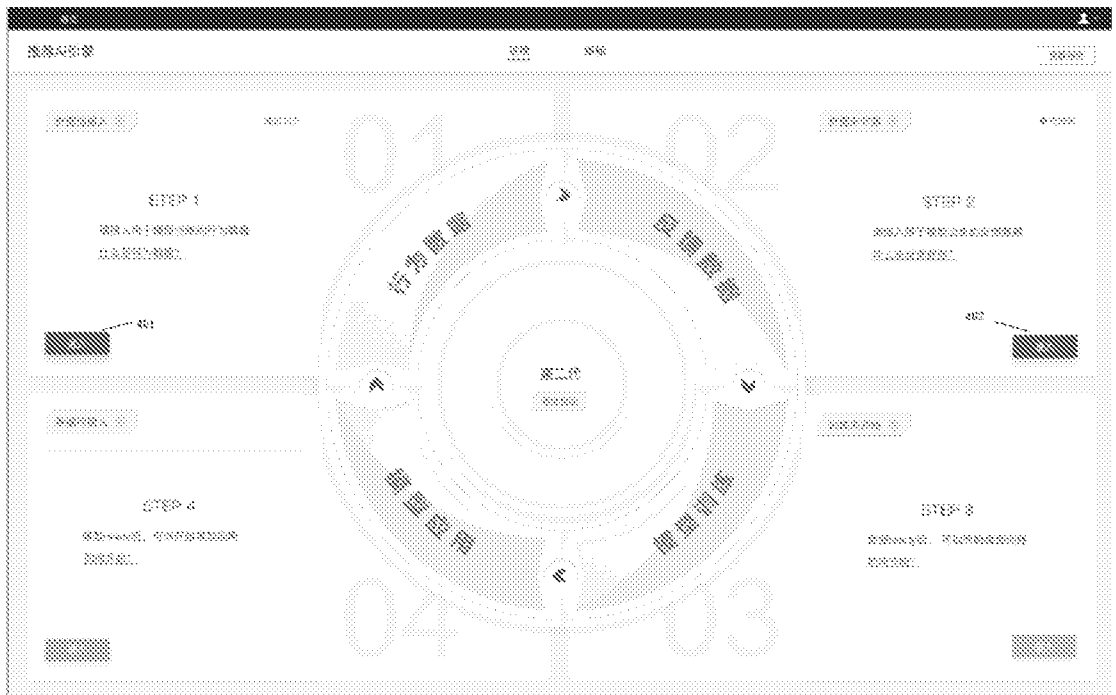


图 3

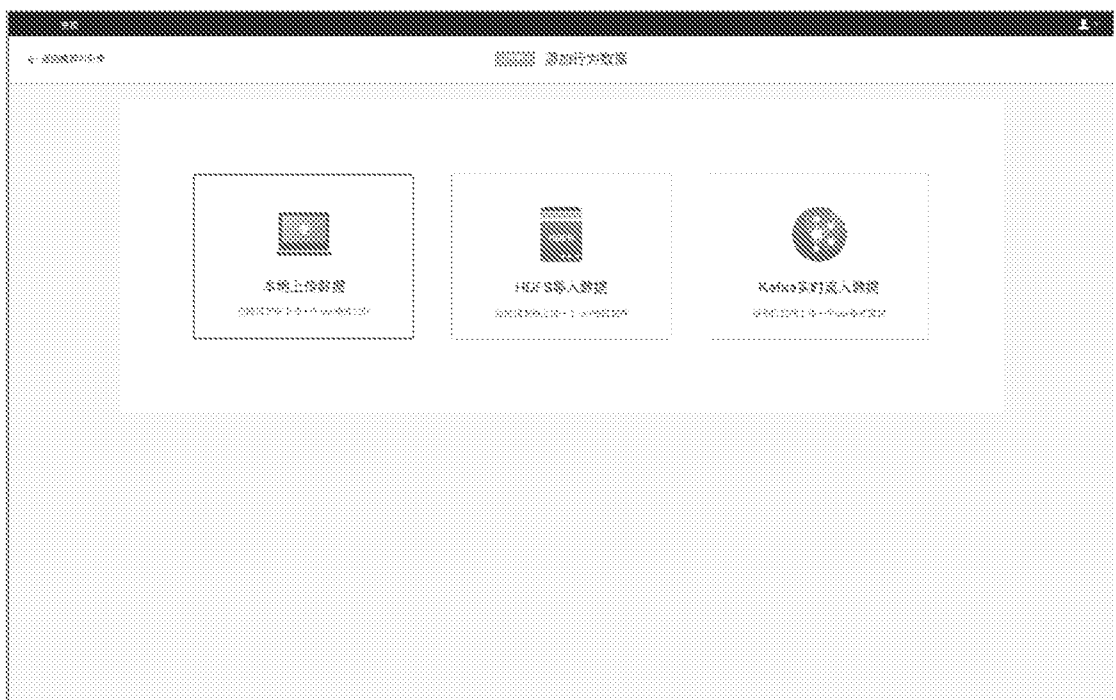


图 4

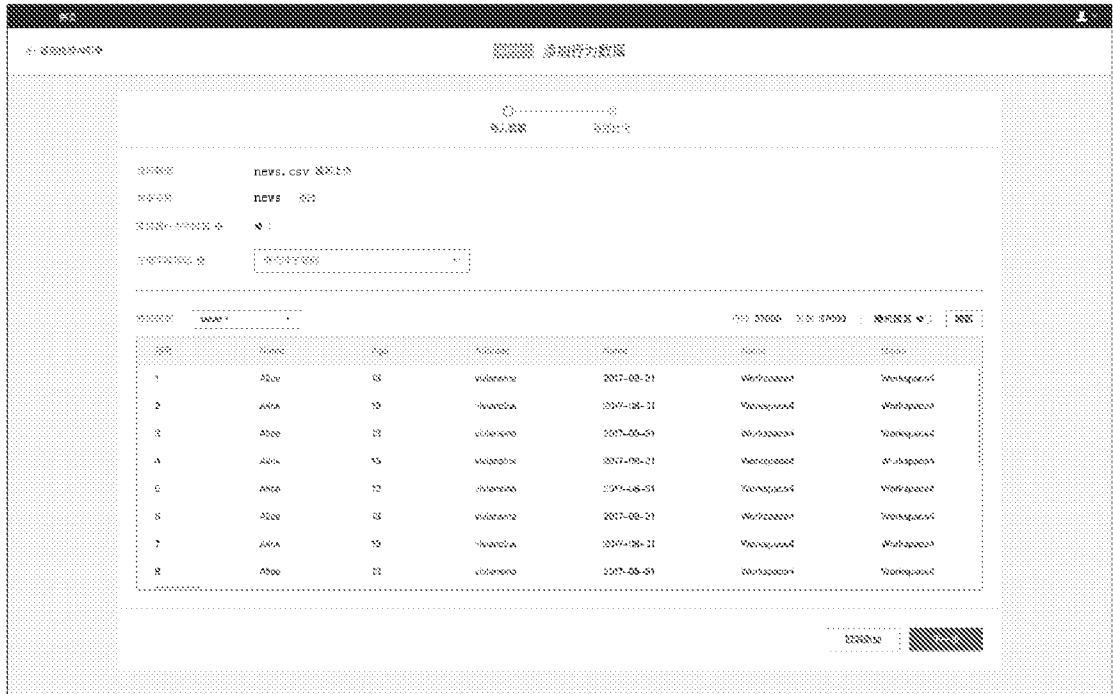


图 5

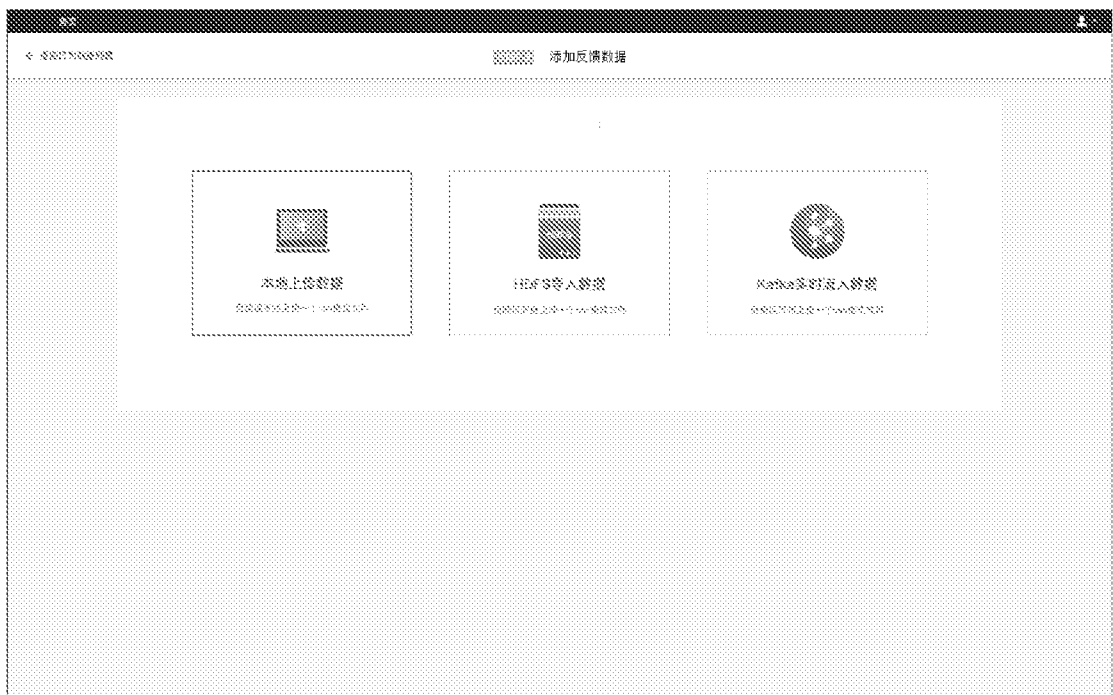


图 6

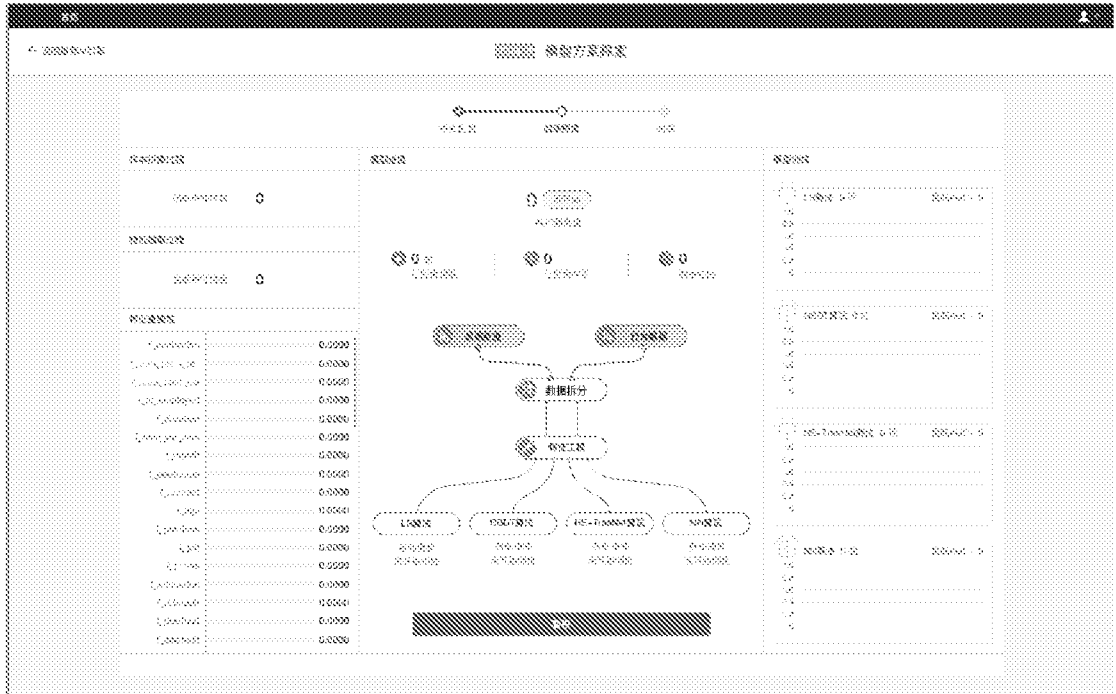


图 7



图 8

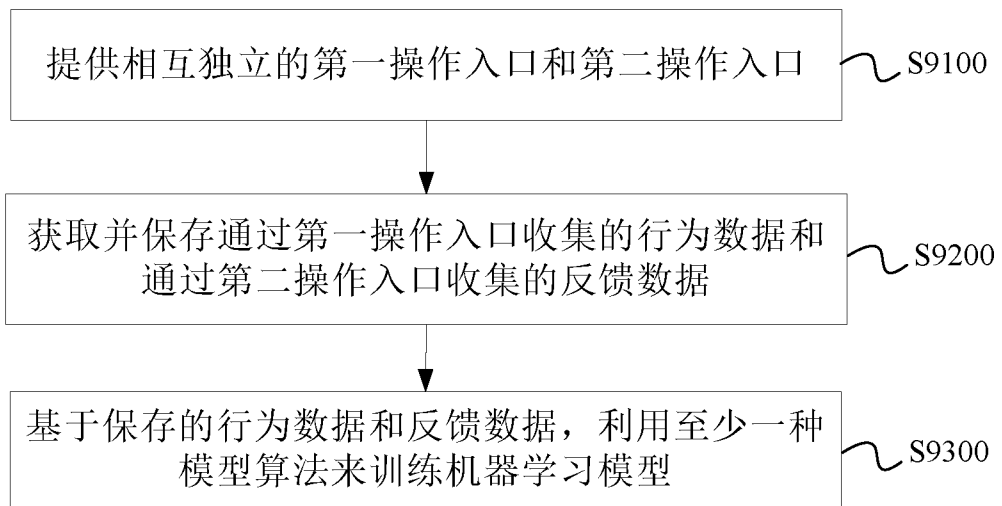


图 9

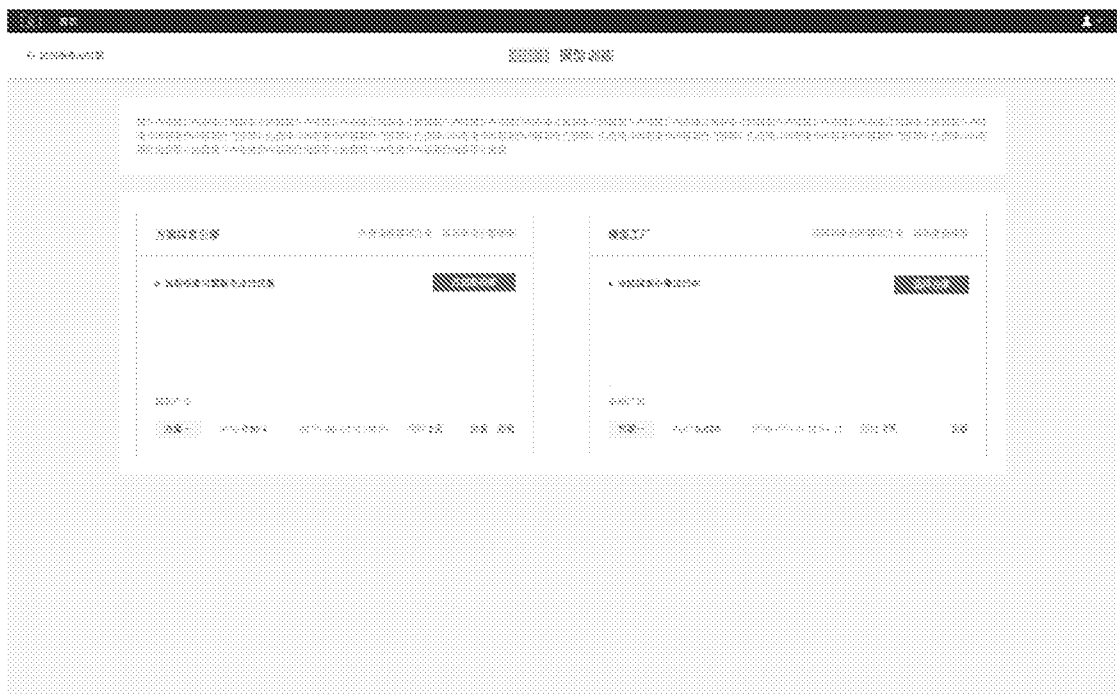


图 10



图 11



图 12



图 16



图 17



图 18

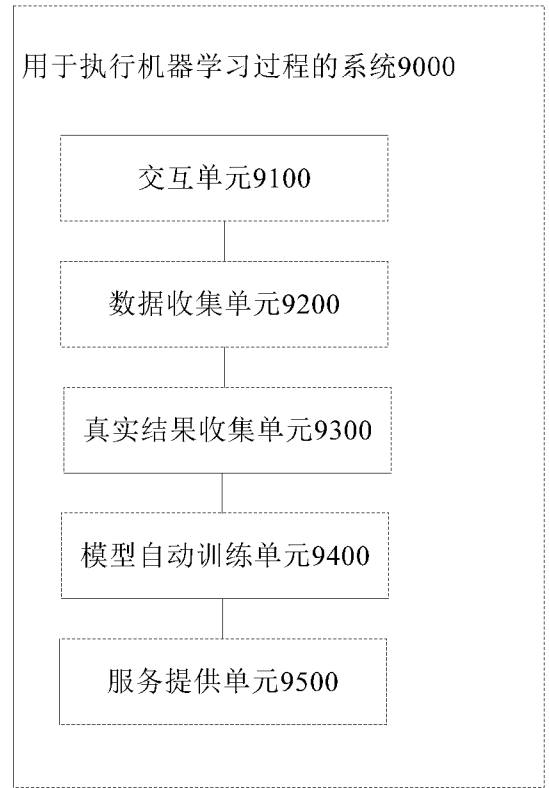


图 19

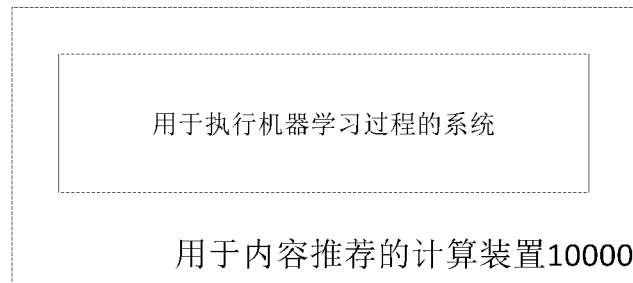


图 20

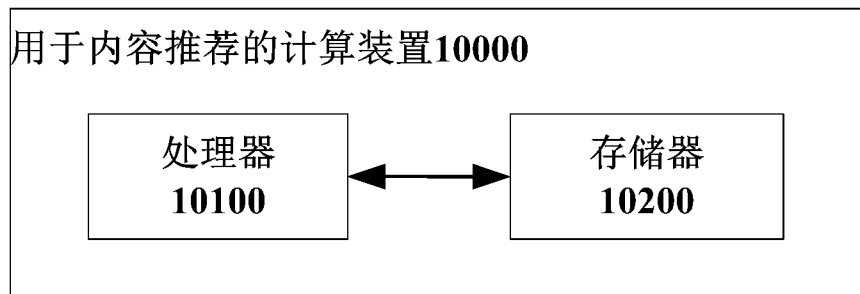


图 21

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/094474

A. CLASSIFICATION OF SUBJECT MATTER		
G06N 5/04(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
G06N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI, EPODOC, CNKI, CNPAT, IEEE: 神经网络, 训练, 结果, 真实, 预测值, 数据, 更新, 模型, network neural, prediction, update, training, result, real, data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 105930934 A (BEIJING WUSI CHUANGXIANG TECHNOLOGY CO., LTD.) 07 September 2016 (2016-09-07) description, paragraphs [0024]-[0054] and [0139]-[0146], and figures 2-6	1-38
Y	US 2015278706 A1 (TELEFONAKTIEBOLAGET L M ERICSSONPUBL) 01 October 2015 (2015-10-01) description, paragraph [0003]	1-38
A	CN 108009643 A (TSINGHUA UNIVERSITY) 08 May 2018 (2018-05-08) entire document	1-38
A	CN 107169573 A (4PARADIGM (BEIJING) TECHNOLOGY CO., LTD.) 15 September 2017 (2017-09-15) entire document	1-38
A	CN 106910013 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 30 June 2017 (2017-06-30) entire document	1-28
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
09 September 2019		26 September 2019
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088 China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/094474

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	105930934	A	07 September 2016	CN	105930934	B	14 August 2018
				WO	2017186048	A1	02 November 2017
				US	2019147350	A1	16 May 2019
				EP	3451250	A1	06 March 2019
				CN	108960514	A	07 December 2018

US	2015278706	A1	01 October 2015	None			

CN	108009643	A	08 May 2018	CN	108009643	B	30 October 2018

CN	107169573	A	15 September 2017	None			

CN	106910013	A	30 June 2017	None			

国际检索报告

国际申请号

PCT/CN2019/094474

<p>A. 主题的分类</p> <p>G06N 5/04 (2006.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																				
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>WPI, EPDOC, CNKI, CNPAT, IEEE: 神经网络, 训练, 结果, 真实, 预测值, 数据, 更新, 模型, network neural, prediction, update, training, result, real, data</p>																				
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 105930934 A (北京物思创想科技有限公司) 2016年 9月 7日 (2016 - 09 - 07) 说明书第[0024]-[0054], [0139]-[0146]段、附图2-6</td> <td>1-38</td> </tr> <tr> <td>Y</td> <td>US 2015278706 A1 (TELEFONAKTIEBOLAGET L M ERICSSONPUBL) 2015年 10月 1日 (2015 - 10 - 01) 说明书第[0003]段</td> <td>1-38</td> </tr> <tr> <td>A</td> <td>CN 108009643 A (清华大学) 2018年 5月 8日 (2018 - 05 - 08) 全文</td> <td>1-38</td> </tr> <tr> <td>A</td> <td>CN 107169573 A (第四范式北京技术有限公司) 2017年 9月 15日 (2017 - 09 - 15) 全文</td> <td>1-38</td> </tr> <tr> <td>A</td> <td>CN 106910013 A (中国科学院自动化研究所) 2017年 6月 30日 (2017 - 06 - 30) 全文</td> <td>1-28</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 105930934 A (北京物思创想科技有限公司) 2016年 9月 7日 (2016 - 09 - 07) 说明书第[0024]-[0054], [0139]-[0146]段、附图2-6	1-38	Y	US 2015278706 A1 (TELEFONAKTIEBOLAGET L M ERICSSONPUBL) 2015年 10月 1日 (2015 - 10 - 01) 说明书第[0003]段	1-38	A	CN 108009643 A (清华大学) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-38	A	CN 107169573 A (第四范式北京技术有限公司) 2017年 9月 15日 (2017 - 09 - 15) 全文	1-38	A	CN 106910013 A (中国科学院自动化研究所) 2017年 6月 30日 (2017 - 06 - 30) 全文	1-28
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
Y	CN 105930934 A (北京物思创想科技有限公司) 2016年 9月 7日 (2016 - 09 - 07) 说明书第[0024]-[0054], [0139]-[0146]段、附图2-6	1-38																		
Y	US 2015278706 A1 (TELEFONAKTIEBOLAGET L M ERICSSONPUBL) 2015年 10月 1日 (2015 - 10 - 01) 说明书第[0003]段	1-38																		
A	CN 108009643 A (清华大学) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-38																		
A	CN 107169573 A (第四范式北京技术有限公司) 2017年 9月 15日 (2017 - 09 - 15) 全文	1-38																		
A	CN 106910013 A (中国科学院自动化研究所) 2017年 6月 30日 (2017 - 06 - 30) 全文	1-28																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2019年 9月 9日</p>		<p>国际检索报告邮寄日期</p> <p>2019年 9月 26日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>张乾楨</p> <p>电话号码 86-(10)-53961401</p>																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/094474

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	105930934	A	2016年 9月 7日	CN	105930934	B	2018年 8月 14日
				WO	2017186048	A1	2017年 11月 2日
				US	2019147350	A1	2019年 5月 16日
				EP	3451250	A1	2019年 3月 6日
				CN	108960514	A	2018年 12月 7日
US	2015278706	A1	2015年 10月 1日	无			
CN	108009643	A	2018年 5月 8日	CN	108009643	B	2018年 10月 30日
CN	107169573	A	2017年 9月 15日	无			
CN	106910013	A	2017年 6月 30日	无			