

机器学习建模过程的实现方法和系统

- 5 本申请要求申请号为 201910525815.4，申请日为 2019 年 6 月 18 日，名称为“机器学习建模过程的实现方法和系统”的中国专利申请的优先权，其中，上述申请公开的内容通过引用结合在本申请中。

技术领域

- 10 本公开总体说来涉及机器学习领域，更具体地说，涉及一种机器学习建模过程的实现方法，机器学习建模过程的实现系统和存储介质。

背景技术

- 15 在传统机器学习建模过程中，需要机器学习专家人工确定所有决策。确定决策的过程往往依赖于人的机器学习经验以及对业务的理解，并且经常需要进行穷举与试错。这一方面需要专门的机器学习人才，另一方面需要大量的人力资源。而在决策空间较大的场景中，通过人工的方式往往无法有效地找到较优的决策。总而言之，传统机器学习建模过程需要专门的人才，消耗大量人力且效率低下。

- 20 自动机器学习（AutoML）技术的目的是使用自动化的数据驱动方式来确定机器学习解决方案。自动机器学习技术虽然能够在一定程度上解决传统机器学习对机器学习人才需求量大、人力资源消耗大的问题，但是现有的自动机器学习技术的搜索效率有限，在计算资源受限、决策空间较大、数据量较大的场景中很难获得较好效果。

25

发明内容

本公开的示例性实施例旨在克服现有的自动机器学习算法的效率低下问题。

- 30 根据本公开的第一个方面，提出了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，指令在被至少一个计算装置运行时，促使至少一个计算装置执行机器学习建模过程的实现方法的以下步骤：

获取用户针对至少部分机器学习建模过程设定的配置信息，所述配置信息用于表征所述至少部分机器学习建模过程的执行策略，所述配置信息包括执行策略已确定的第一部分和执行策略待确定的第二部分；

- 35 利用自动机器学习技术确定所述第二部分的执行策略；

基于所述第一部分的执行策略和所确定的所述第二部分的执行策略，执行所述至少部分机器学习建模过程，得到与所述至少部分机器学习建模过程对应的结果。

根据本公开的第二个方面，还提出了一种由至少一个计算装置执行的机

器学习建模过程的实现方法，包括：获取用户针对至少部分机器学习建模过程设定的配置信息，配置信息用于表征至少部分机器学习建模过程的执行策略，配置信息包括执行策略已确定的第一部分和执行策略待确定的第二部分；利用自动机器学习技术确定第二部分的执行策略；基于第一部分的执行策略和所确定的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

根据本公开的第三个方面，还提出了一种机器学习建模过程的实现系统，包括：交互装置，用于获取用户针对至少部分机器学习建模过程设定的配置信息，配置信息用于表征至少部分机器学习建模过程的执行策略，配置信息包括执行策略已确定的第一部分和执行策略待确定的第二部分；机器学习处理装置，用于利用自动机器学习技术确定第二部分的执行策略，以及基于第一部分的执行策略和所确定的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

根据本公开的第四个方面，还提出了一种存储指令的计算机可读存储介质，其中，当指令被至少一个计算装置运行时，促使至少一个计算装置执行如本公开第二个方面述及的方法。

根据本公开示例性实施例的机器学习建模过程的实现方法和系统，通过参考用户的机器学习经验以及对业务的理解，从用户设定的配置信息中获取用户能够确定执行策略的第一部分；对于用户无法确定执行策略的第二部分，再利用自动机器学习技术确定该部分的执行策略；从而可以将人与机器的优势结合起来，提高机器学习过程的实现效率。

附图说明

从下面结合附图对本公开实施例的详细描述中，本公开的这些和/或其他方面和优点将变得更加清楚并更容易理解，其中：

图 1 示出了根据本公开示例性实施例的机器学习建模过程的实现方法的流程图；

图 2 示出了根据本公开示例性实施例的机器学习开发系统的架构图；

图 3 示出了根据本公开示例性实施例的原始的本体模型的生成过程示意图；

图 4 示出了通过对图 3 所示的本体模型执行操作实现的机器学习建模过程的一个解决方案的示意性流程图；

图 5 示出了基于图 4 构建的计算图的示意图；

图 6 示出了基于图 4 构建的另一种计算图的示意图；

图 7 示出了根据本公开示例性实施例的机器学习建模过程的实现系统的结构框图。

具体实施方式

为了使本领域技术人员更好地理解本公开，下面结合附图和具体实施方式

式对本公开的示例性实施例作进一步详细说明。在此需要说明的是，在本公开中出现的“若干项之中的至少一项”均表示包含“该若干项中的任意一项”、“该若干项中的任意多项的组合”、“该若干项的全体”这三类并列的情况。在本公开中出现的“和/或”、“并且/或者”均表示被其连接的前后两项中的至少一项。例如，“包括 A 和 B 之中的至少一个”、“A 和/或 B”、“A 并且/或者 B”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，“执行步骤一和步骤二之中的至少一个”、“执行步骤一和/或步骤二”、“执行步骤一并且/或者步骤二”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。图 1 示出了根据本公开示例性实施例的机器学习建模过程的实现方法的流程图。其中，图 1 所示的方法可完全通过计算机程序以软件方式实现，还可通过特定配置的计算装置来执行图 1 所示的方法。

需要说明的是，本公开实施例中的所述的机器学习模型可被用于预测图像类别、文本类别、语音情感、欺诈交易、广告点击率等。所述机器学习模型旨在针对相关场景中的对象或事件有关的问题进行预测。例如，可用于预测图像类别、预测图像中文字、预测文本类别、预测语音情感类别、预测欺诈交易、预测广告点击率、预测商品价格等等，使得预测结果可直接作为决策依据或进一步结合其他规则而成为决策依据。

更进一步来说，本公开实施例中的机器学习模型可被用于的场景包括但不限于以下场景：

图像处理场景，包括：光学字符识别 OCR、人脸识别、物体识别和图片分类；更具体地举例来说，OCR 可应用于票据（如发票）识别、手写字识别等，人脸识别可应用安防等领域，物体识别可应用于自动驾驶场景中的交通标志识别，图片分类可应用于电商平台的“拍照购”、“找同款”等。

语音识别场景，包括可通过语音进行人机交互的产品，如手机的语音助手（如苹果手机的 Siri）、智能音箱等；

自然语言处理场景，包括：审查文本（如合同、法律文书和客服记录等）、垃圾内容识别（如垃圾短信识别）和文本分类（情感、意图和主题等）；

自动控制场景，包括：矿井组调节操作预测、风力发电机组调节操作预测和空调系统调节操作预测；具体的对于矿井组可预测开采率高的一组调节操作，对于风力发电机组可预测发电效率高的一组调节操作，对于空调系统，可以预测满足需求的同时节省能耗的一组调节操作；

智能问答场景，包括：聊天机器人和智能客服；

业务决策场景，包括：金融科技领域、医疗领域和市政领域的场景，其中：

金融科技领域包括：营销（如优惠券使用预测、广告点击行为预测、用户画像挖掘等）与获客、反欺诈、反洗钱、承保和信用评级、商品价格预测；

医疗领域包括：疾病筛查和预防、个性化健康管理和辅助诊断；

市政领域包括：社会治理与监管执法、资源环境和设施管理、产业发展

和经济分析、公众服务和民生保障、智慧城市（公交、网约车、共享单车等各类城市资源的调配和管理）；

推荐业务场景，包括：新闻、广告、音乐、咨询、视频和金融产品（如理财、保险等）的推荐；

5 搜索场景，包括：网页搜索、图像搜索、文本搜索、视频搜索等；

异常行为检测场景，包括：国家电网客户用电异常行为检测、网络恶意流量检测、操作日志中的异常行为检测等。

参见图 1，在步骤 S110，获取用户针对至少部分机器学习建模过程设置的配置信息。

10 至少部分机器学习建模过程可以是完整的机器学习建模过程，也可以是机器学习建模过程中的部分流程。作为示例，至少部分机器学习建模过程可以包括但不限于数据拼接过程、数据拆分过程、特征生成过程、模型建立过程、模型训练过程、模型测试过程、模型评估过程、模型应用过程等诸多过程中的一项或多项。

15 配置信息用于表征至少部分机器学习建模过程的执行策略，配置信息包括执行策略已确定的第一部分和执行策略待确定的第二部分。第一部分是指用户基于自身的机器学习经验以及对业务的理解，能够确定执行策略的部分，第二部分是指用户自身无法确定、需要借助机器确定执行策略的部分。

20 举例来说，假设至少部分机器学习建模过程包括特征生成过程和模型建立过程，用户知道选取何种特征工程函数能够得到较好的特征，但不知道选取何种模型结构，那么第一部分可以是指执行策略已确定的特征生成过程，第二部分可以是指执行策略待确定的模型建立过程。

25 作为示例，第一部分可以包括一个或多个执行步骤（为了便于区分，可以称为第一执行步骤），第一执行步骤包括第一输入、第一处理函数以及第一输出，并且/或者第二部分可以包括一个或多个执行步骤（为了便于区分，可以称为第二执行步骤），第二执行步骤包括第二输入、第二处理函数以及第二输出，第二输入、第二处理函数以及第二输出中至少一个待确定。其中，不同的执行步骤（第一执行步骤/第二执行步骤）间具有预定的执行顺序。作为示例，可以在配置信息中标明各个执行步骤的执行顺序。第一处理函数和/或
30 第二处理函数可以包括但不限于数据处理函数、特征工程函数、建模函数、自动机器学习函数、共享与重用函数、版本管理函数、模型应用函数中的一项或多项。

在此需要说明的是，上述内容还可以表述为，本公开还包括如下两项内容中的至少一项：

35 第一部分可以包括一个或多个执行步骤（为了便于区分，可以称为第一执行步骤），第一执行步骤包括第一输入、第一处理函数以及第一输出。

第二部分可以包括一个或多个执行步骤（为了便于区分，可以称为第二执行步骤），第二执行步骤包括第二输入、第二处理函数以及第二输出，第二输入、第二处理函数以及第二输出中至少一个待确定。

在此还需要说明的是，本公开中的第一处理函数和/或第二处理函数，还可以表述为第一处理函数和第二处理函数之中的至少一项。

对于第二部分，用户可以基于对业务的理解以及根据以往的建模经验指定一个决策空间，由机器在用户指定的决策空间内搜索较优的执行策略；或者用户也可以不指定决策空间，而是由机器在可行的决策空间内搜索较优的执行策略。

本公开述及的配置信息可以是指多种形式的信息，也即用户可以通过多种方式设定配置信息。

作为本公开的一个示例，配置信息可以是用户基于预定规则的语言（为了便于区分，可以称为“第二规则语言”）编写的。其中，第二规则语言可以是指符合预定规则形式的语言。例如，第二规则语言可以是指领域专用语言（Domain Specific Language，简称 DSL）。关于领域专用语言可以参见下文相关描述，此处暂不赘述。

作为本公开的另一个示例，配置信息也可以是用户通过对向用户展示的用于对至少部分机器学习建模过程进行设置的交互界面执行的操作生成的。换言之，用户可以通过可视化操作（如拖拽）的方式来设定配置信息。例如，用户可以通过在交互界面中搭建用于表征至少部分机器信息建模过程的执行策略的流程图的方式，来设定配置信息，也即配置信息可以是用户在交互界面中搭建的流程图。

另外配置信息也可以是指用户通过其他方式设定的，对此本公开不再赘述。

在步骤 S120，利用自动机器学习技术确定第二部分的执行策略。

可以利用自动机器学习技术在第二部分对应的决策空间内搜索较优的执行策略，以确定第二部分的执行策略。其中，在用户指定了第二部分的决策空间的情况下，可以利用自动机器学习技术在用户指定的决策空间内搜索较优的执行策略，作为第二部分的执行策略。在用户未指定第二部分的决策空间的情况下，可以利用自动机器学习技术在可行的决策空间内搜索较优的执行策略，作为第二部分的执行策略。

在本公开中，可以利用优化算法、搜索算法等多种算法在第二部分的决策空间内确定一个较优的执行策略。其中，决策空间用于表征第二部分的执行策略的取值集合，该决策空间可以由用户指定的，也可以是由机器确定的。例如，可以对第二部分的决策空间中多个取值不同的执行策略进行评测，然后将评测结果较好的执行策略确定为第二部分的执行策略。

作为示例，可以按照第一部分的执行策略以及赋值后的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果；对结果进行评测；根据评测结果重新为第二部分的执行策略赋值；重复上述流程，直到满足预定条件，得到第二部分的多个取值不同的执行策略的评测结果，然后将评测结果较好的执行策略确定为第二部分的执行策略。此处述及的预定条件可以是预先设定的标准，如可以是执行次数

或者评测结果达到预定阈值。另外此处述及的预定条件也可以是由用户做出的终止操作，也即可以由用户终止迭代过程。

在确定了第二部分的执行策略后，至少部分机器学习建模过程的整个执行流程也就确定了。因此在确定了第二部分的执行策略后，可以执行步骤 S130，
5 基于第一部分的执行策略和所确定的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

综上，针对现有自动机器学习技术的效率低下问题，本公开可以参考用户的机器学习经验以及对业务的理解，从用户设定的配置信息中获取用户能够确定执行策略的第一部分。对于用户无法确定执行策略的第二部分，再利
10 用自动机器学习技术确定该部分的执行策略。由此，本公开可以将人与机器的优势结合起来，从而可以提高机器学习过程的实现效率。

并且现有的自动机器学习技术大多以黑盒 (blackbox) 的形式完成一项或多项机器学习功能，用户难以进行干涉。在本公开中，用户可以根据自身的机器学习经验以及对业务的理解，灵活地确定哪些是由人工可以确定的执行
15 策略，哪些是需要机器确定的执行策略。由此，本公开可以实现从完全自动化 (即全部执行策略均由机器确定) 到完全人工 (即用户可以确定全部执行策略)。

根据上文描述可知，用户设定的配置信息一般不能被后端引擎识别，为此本公开提出，可以将配置信息翻译为后端引擎对应的语言或指令，以便
20 后端引擎能够基于翻译后的语言或指令执行步骤 S120 和/或步骤 S130。在此需要说明的是，本公开中的执行步骤 S120 和/或步骤 S130，还可以表述为执行步骤 S120 和步骤 S130 之中的至少一项。

作为示例，后端引擎可以包括计算图引擎和执行引擎，执行引擎可以包括但不限于优化引擎、机器学习引擎以及自动机器学习引擎中的一个或多个，
25 并且不同的引擎对应的语言类型可以相同，也可以不同。

可以由计算图引擎基于语言或指令生成第一计算图，第一计算图用于表征确定第二部分的执行策略的实现流程，可以由计算图引擎通过调用执行引擎执行第一计算图，确定第二部分的执行策略。

还可以由计算图引擎基于语言或指令生成第二计算图，第二计算图用于表征至少部分机器学习建模过程的实现流程，可以由计算图引擎通过调用执行引擎执行第二计算图，得到与至少部分机器学习建模过程对应的结果。
30

在现有技术中，大部分自动机器学习算法只能考虑具有严谨定义的指标 (例如可以用显示、明确的方法计算的指标)，但是实际业务中很多重要指标 (如是否发生数据泄露、结果的可解释性) 难以给出严谨定义。导致自动机器学习算法无法考虑这些指标，难以产生符合实际业务需求的结果。并且现
35 有技术大多以黑盒 (blackbox) 的形式完成一项或多项机器学习功能，用户难以进行干涉。

为了解决该问题，本公开提出，在执行步骤 S120 和/或步骤 S130 的过程中，还可以实时向用户输出当前执行结果；并且/或者在执行步骤 S120 和/或

步骤 S130 的过程中,还可以获取用户针对配置信息做出的调整,并基于调整后的配置信息执行至少部分机器学习建模过程。

在此需要说明的是,上述内容还可以表述为,本公开还可以包括以下步骤中的至少一个步骤:

- 5 在执行步骤 S120 和/或步骤 S130 的过程中,实时向用户输出当前执行结果;

在执行步骤 S120 和/或步骤 S130 的过程中,获取用户针对配置信息做出的调整,并基于调整后的配置信息执行至少部分机器学习建模过程。

- 10 由此,本公开可以允许用户随时查看执行结果,便于用户基于自己对业务的理解判断当前的方案能否得到符合业务需求的结果,而不强求用户给出目标的严谨定义,从而可以提高机器学习建模过程的可控性。

- 15 现有的自动机器学习算法是基于具体数据来确定解决方案的,在将利用自动机器学习算法确定的解决方案应用到具体场景中时,由于具体场景中的数据 and 确定解决方案时所使用的数据的存储形式可能不一致,使得还需要根据数据的实际存储形式对解决方案进行修改,实现起来较为繁琐。

- 20 为了解决该问题,本公开提出,可以获取数据描述信息,数据描述信息描述了一种或多种数据、数据的属性以及数据和/或属性的关系,基于数据描述信息,得到用于结构化地表征数据、属性以及关系的数据模型。在此需要说明的是,本公开中的数据和/或属性,可以表述为数据和属性之中的至少一项。

- 25 此处述及的数据表征的是一种数据实体,如可以是交易数据、用户数据、物品数据等等。属性用于描述数据的特性,可以是指数据中的一个字段,例如一条用户数据可以包括年龄、性别、职业等多个字段,因此用户数据可以包括但不限于年龄、性别、职业等多个属性。数据描述信息中针对属性的描述可以包括该属性是从何处获取的,也即属性所对应的实际数据的存储方式,例如某个属性对应于哪个文件中的哪一列。数据和/或属性的关系包括数据与数据的关系、数据与属性的关系以及属性与属性的关系。

- 30 需要说明的是,在本公开中,数据与属性之间并无严格界限,属性本身也可以认为是一种数据,如何看待属性取决于实际应用中是强调其作为数据的属性,还是作为数据。

- 35 在本公开中,数据描述信息可以是用户基于预定规则的语言(为了便于区分,可以称为“第一规则语言”)编写的。关于第一规则语言的具体形式,本公开不做限定。根据上文描述可知,数据描述信息中包含了数据和/或属性间的逻辑关系以及实际数据的存储方式,因此通过对数据描述信息进行解析,就可以得到能够结构化地表征数据、属性以及关系的数据模型,其中数据模型中还可以包括数据的实际存储位置。作为示例,可以使用能够识别第一规则语言的读取函数读取数据描述信息,并基于数据描述信息得到数据模型。

配置信息可以是用户根据数据模型设定的、用于表征基于数据模型所表征的数据和/或属性实现至少部分机器学习建模过程的执行策略。例如,可以

向用户展示数据模型，以便用户根据数据模型设定配置信息。

可选地，还可以输出建议信息，建议信息用于表征针对数据模型的可行操作。由此，用户在设定配置信息时可以参考向其展示的建议信息。例如，可以根据数据模型（主要是数据及其关系）给出一个包含全部或部分可行操作的集合（此处述及的可行操作是指针对当前数据模型的合法的、具体的操作，例如：以某些数据和变量/常量为输入，调用某个特定的特征工程函数，生成一个或一组属性）作为建议策略。用户可以查看、修改、采纳该建议策略中的全部或部分操作。

通过将数据模型作为用户与实际数据间的接口，在底层完成不同存储形式的数据向数据模型的转换，从而可以让用户设定配置信息时，可以专注于数据的含义以及业务逻辑而不需要考虑其实际的存储形式，减少方案部署时因数据存储形式的不一致而对方案的修改。

也就是说，在数据的存储形式发生变化时，只需要修改数据模型描述信息，导入新的数据模型描述信息就可以完成针对这些改变的修改，而不需要修改配置信息。或者在使用的数据的存储形式发生变化时，可以直接修改数据模型中发生变化的部分，在这种情况下，也不需要修改配置信息。

为了更好地体现数据模型的效果，下面就基于数据模型设定配置信息时，步骤 S120、步骤 S130 的实现过程进行示例性说明。

作为示例，步骤 S120 可以包括：步骤 A1、基于数据模型，获取执行至少部分机器学习建模过程所需的数据；步骤 A2、为第二部分的执行策略赋值；步骤 A3、按照第一部分的执行策略以及赋值后的第二部分的执行策略，针对数据执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果；步骤 A4、对结果进行评测。

由此可以得到第二部分的赋值后的执行策略的评测结果。基于步骤 A1 至步骤 A4 的评测流程，可以利用优化算法、搜索算法等多种算法在第二部分的决策空间内确定一个较优的执行策略。例如，可以基于上述步骤 A1 至步骤 A4，进一步得到第二部分的多个取值不同的执行策略的评测结果，然后将评测结果较好的执行策略确定为第二部分的执行策略。可选地，基于上述步骤 A1 至步骤 A4，得到所述第二部分的多个取值不同的执行策略的评测结果，将评测结果较好的执行策略确定为所述第二部分的执行策略的步骤可以包括：步骤 A5、根据评测结果重新为第二部分的执行策略赋值；重复上述步骤 A3 至步骤 A5，直到满足预定条件，将评测结果较好的执行策略确定为第二部分的执行策略。

作为示例，步骤 S130 可以包括：基于数据模型，获取执行至少部分机器学习建模过程所需的数据；基于第一部分的执行策略和所确定的第二部分的执行策略，针对数据执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

可以看到，步骤 S120、步骤 S130 的执行过程中，是根据数据模型获取实际数据的。因此，在实际数据的存储位置发生变化时，仅需要更新数据模

型，无需修改步骤 S120、步骤 S130 的执行策略。由此，在利用本公开确定了机器学习建模过程的解决方案后，将方案部署到线上时，即使数据存储形式发生了变化，也无需对方案本身进行修改。

如上文所述，可以基于领域专用语言编写配置信息。下面就基于领域专用语言编写配置信息时，本公开的实现过程进行示例性说明。

本公开述及的领域专用语言是指专注于机器学习领域的语言，该语言定义了多种实体和一系列函数（functions）。

实体用于表征机器学习领域中具有预定含义的对象，并且/或者实体能够执行预定的动作。在此需要说明的是，上述内容还可以表述为：实体被配置为用于表征机器学习领域中具有预定含义的对象和用于执行预定的动作之中的至少一个。

例如实体可以包括但不限于数据实体、属性实体、常量实体、变量实体、模型实体、决策实体、目标实体。其中，数据实体用于表征机器学习语境下的数据，能够执行获取实际数据的动作；属性实体用于表征数据实体的属性，能够执行获取属性所对应的实际数据的动作；常量实体用于表征常量，存储了常量的值；变量实体用于表征变量，存储了变量的值；模型实体用于表征机器学习语境下的模型，能够执行训练模型和/或使用模型进行预测的动作；决策实体用于表征期望由机器确定的决策，能够执行生成决策和/或获取最优决策的动作；目标实体用于表征评测结果，目标实体能够执行优化和/或计算评测结果的动作。

在此需要说明的是，本公开中的执行训练模型和/或使用模型进行预测的动作，还可以表述为执行训练模型和使用模型进行预测之中的至少一个的动作。

本公开中的执行生成决策和/或获取最优决策的动作，还可以表述为执行生成决策和获取最优决策之中的至少一个的动作。

本公开中的执行优化和/或计算评测结果的动作，还可以表述为执行优化和计算评测结果之中的至少一个的动作。

一系列函数是指能够实现预定功能的函数，用户可以通过调用特定的函数实现特定的功能。作为示例，一系列函数能够实现的功能可以包括但不限于数据处理、特征工程、建模、模型应用、交互式机器学习与自动机器学习解决方案开发、自动机器学习技术的应用、本体模型的创建、修改、重用与版本维护。

如上文，第一部分包括一个或多个第一执行步骤，第一执行步骤包括第一输入、第一处理函数以及第一输出，并且/或者第二部分包括一个或多个第二执行步骤，第二执行步骤包括第二输入、第二处理函数以及第二输出，第二输入、第二处理函数以及第二输出中至少一个待确定。

因此，在基于领域专用语言编写配置信息时，第一输入可以包括一个或多个实体（为了便于区分，可以称为第一实体），第一实体用于表征机器学习领域中具有预定含义的对象，并且/或者第一实体能够执行预定的动作。第二

输入可以包括一个或多个实体（为了便于区分，可以称为第二实体），第二实体用于表征机器学习领域中具有预定含义的对象，并且/或者第二实体能够执行预定的动作。

- 在此需要说明的是，上述内容还可以表述为：第一输入包括一个或多个
- 5 第一实体，所述第一实体被配置为用于表征机器学习领域中具有预定含义的对象和用于执行预定的动作之中的至少一个；第二输入包括一个或多个第二实体，第二实体被配置为用于表征机器学习领域中具有预定含义的对象用于执行预定的动作之中的至少一个。

- 10 第一实体和/或第二实体可以包括但不限于数据实体、属性实体、常量实体、变量实体、模型实体中的一种或多种。在此需要说明的是，本公开中的第一实体和/或第二实体，还可以表述为第一实体和第二实体之中的至少一个。

- 第二执行步骤是指执行策略待确定的步骤，因此第二实体还可以包括决策实体。也就是说，可以利用决策实体来标识出需要由机器确定执行策略的部分。可选地，决策实体可以包括决策的取值范围，以便机器在该取值范围
- 15 内确定决策实体的最优决策。

- 配置信息还可以包括执行策略已确定的优化步骤，优化步骤的输入包括至少部分机器学习建模过程的结果，优化步骤的处理函数为目标定义函数，目标定义函数用于对结果进行评测，以得到目标实体。其中，目标实体可以是用户从已有的多个目标实体中选择的
- 20 目标实体，也可以是由用户定义的目标实体。

- 由此，步骤 S120 包括：执行优化步骤，其中，执行优化步骤的流程包括：步骤 B1、对执行至少部分机器学习建模过程所需的原始数据实体和/或原始属性实体，执行获取实际数据动作，以得到对应的实际数据；步骤 B2、对决策实体执行生成决策动作，以得到决策实体的取值；步骤 B3、按照第一部分的
- 25 执行策略以及赋值后的第二部分的执行策略，针对数据执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果；步骤 B4、对目标实体执行计算评测结果的动作。在此需要说明的是，本公开中的原始数据实体和/或原始属性实体，还可以表述为原始数据实体和原始属性实体之中的至少一个。

- 30 由此可以得到决策实体赋值后评测结果。基于步骤 B1 至步骤 B4，可以利用优化算法、搜索算法等多种算法在决策实体的取值空间内确定一个较优的取值。例如可以基于上述步骤 B1 至步骤 B4，得到决策实体的多个取值的评测结果，将评测结果较好的取值确定为决策实体的最优决策。其中，原始数据实体、原始属性实体是指原本存在的，而非是经过函数处理生成的。

- 35 作为示例，基于上述步骤 B1 至步骤 B4，得到决策实体的多个取值的评测结果，将评测结果较好的取值确定为决策实体的最优决策的步骤可以包括：步骤 B5、基于评测结果确定决策实体的下一轮取值；重复上述步骤 B2 至步骤 B5，直到满足预定条件，将评测结果较好的取值确定为决策实体的最优决策。

步骤 S130 包括：对执行至少部分机器学习建模过程所需的原始数据实体和/或原始属性实体，执行获取实际数据动作，以得到对应的实际数据；获取决策实体的最优决策；按照第一部分的执行策略和所确定的第二部分的执行策略，针对实际数据执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

可选地，在整个执行过程中，可以向用户展示第一实体和/或第二实体；并且/或者接收用户针对第一实体和/或第二实体的修改。由此，可以使得机器学习建模过程可控。

在此需要说明的是，上述内容还可以表述为，在整个执行过程中还包括以下步骤中的至少一个步骤：

向用户展示所述第一实体和第二实体之中的至少一个；

接收用户针对第一实体和第二实体之中的至少一个的修改。

应用例

图 2 示出了根据本公开示例性实施例的机器学习开发系统的架构图。可以由图 2 所示的机器学习开发系统执行本公开的机器学习建模过程的实现方法。

如图 2 所示，机器学习开发系统自上往下依次是交互式开发界面、领域专用语言、解释器、后端引擎以及数据管理系统。其中，越上层越容易被用户接触。

交互式开发界面是用户与机器交互，共同开发机器学习或自动机器学习解决方案的界面。交互式开发界面的功能可以包含但不限于：本体模型的可视化功能；机器学习解决方案开发过程中的决策建议功能（向用户建议可行的操作、可能较优的决策等）；后台运行代码及随时查看运行进度与结果（或中间结果）的功能。

领域专用语言是指专注于机器学习领域的语言，是用户与机器交互用户用于开发、修改、维护、重用解决方案的编程语言。

如图 2 所示，领域专用语言主要包括本体模型和一系列函数。其中，图 2 示出的本体模型是指基于领域专用语言实现的针对至少部分机器学习建模过程的机器学习解决方案。图 2 中示出的本体模型主要由实体、关系以及属性构成。该本体模型可以既作为用户与后端引擎之间的交互媒介，也可以作为用户进行开发迭代、版本管理、分享以及重用的载体。关于实体、属性、关系以及一系列函数可以参见上文相关描述，此处不再赘述。

解释器主要用于将领域专用语言解释为不同后端引擎的语言或指令，如可执行代码。

后端引擎包含部分现有（包含自研的）的机器学习引擎、自动机器学习引擎、优化引擎、计算图引擎，并在未来可以接入其他引擎。

数据管理系统可以用于实现数据的存储、访问以及转移。

图 3 示出了根据本公开示例性实施例的原始的本体模型的生成过程示意图。图 3 示出的本体模型，也即上文述及的数据模型，图 3 中示出的本体模

型描述文件也即上文述及的数据模型描述文件。

参见图 3, 可以由与本公开述及的领域专用语言对应的读取函数, 通过读取符合特定格式的本体模型描述文件, 得到原始的本体模型。该描述性文件描述了本体中的实体、属性以及关系。此处尤其需要注意属性描述。属性描述声明了每一个实体的每一个属性所对应的实际数据的存储位置 (例如, 某个文件的某一行)。这些信息会记录在本体模型的每一个属性中。当全部和部分数据的存储位置或形式发生改变时, 用户只需要修改描述文件 (或重载一个新的描述性文件), 就可以完成针对这些改变 (数据的存储位置或形式发生改变) 的修改, 而不需要修改本体模型。另一种可选方法是直接修改本体模型中属性所存储的信息, 在这种情况下, 不需要修改本体模型的其他部分。

图 4 示出了通过对图 3 所示的本体模型执行操作实现的机器学习建模过程的一个解决方案的示意性流程图。其中图 4 中本体模型 (图 3 所示的本体模型) 之外的部分均可以视为用户根据本体模型设定的配置信息。也就是说, 原始的本体模型和用户设定的配置信息构成了一个完整的机器学习解决方案, 例如如图 4 所示的解决方案。其中, 图 4 所示的解决方案可以视为是基于领域专用语言构造的完整的本体模型, 该本体模型是指用于表征完整的机器学习解决方案, 不同于图 3 所示的仅用于表示数据结构的本体模型 (也即数据模型)。

图 4 中的矩形表示实体, 圆形标识属性、菱形表示函数, 箭头连线表示关系, 圆头连线链接实体与对应的属性。实线的形状以及连线表示原始的本体模型 (也即图 3 所示的本体模型), 虚线的形状以及连线表示后续操作生成的实体、关系和属性。数字表示操作的顺序, 例如标数值 1 的为一个特征工程函数, 以及它的三个输入和一个输出, 这表示对本体模型的第一个操作是执行该特征工程函数, 并且该函数有三个输入和一个输出。

如图 4 所示, 在本实施例中, 原始的本体模型涉及两种关系: 1) 普通的关系, 例如关系 1、关系 2、关系 3 等; 2) 由函数构成的关系, 如由数字 1 表示出的: “属性实体 3 是由特征工程函数, 以数据实体 3、数据实体 4、常量/变量实体 1 和决策实体 1 为输入生成的输出” 这一关系。

图 4 中的 1、2、3、4 这四个步骤依次对原始的本体模型进行了修改。可以暂时忽略这些操作涉及的决策实体, 将其简单认为是某个实体即可, 同样可以暂时忽略步骤 5、6, 将在下文详述。这里将分步解释每个步骤的效果:

步骤 1: 以数据实体 3、数据实体 4、常量/变量实体 1、决策实体 1 为输入, 调用特征工程函数, 生成一个作为输出的属性实体 3;

步骤 2: 以属性实体 3 为输入, 调用增加属性函数, 修改数据实体 1;

步骤 3: 以步骤 2 之后的数据实体 1、变量/常量实体 2 为输入, 调用数据划分函数, 输出数据实体 5 和数据实体 6;

步骤 4: 以数据实体 5、决策实体 2 为输入, 调用建模函数, 生成一个作为输出的模型实体 1。

图 4 中的决策实体 1 和决策实体 2 代表由用户指定的, 希望交由机器决

定的决策。在指定的时候，决策实体处于“待定”状态。为了让机器能够做出正确的决策，用户还需要告知机器一个目标，令其以优化这个目标为目的，确定每一个决策。为了描述优化目标并执行优化，用户设定了 5、6 两个操作（对应于上文述及的优化步骤）。

5 步骤 5：以模型实体 1、数据实体 6 为输入，调用目标定义函数，输出目标实体 1（代表，例如，模型的精度）；

 步骤 6：令目标实体 1 执行调用优化函数，作出针对决策实体 1 和决策实体 2 的决策。

10 在执行步骤 6 后，所有与目标实体 1 相关（下文详述）的决策实体将获得一个确定的值（对应机器作出的决策），并且其状态变为“确定”。

 在确定了决策实体的取值后，就可以执行步骤 7，训练模型。

 下面就图 4 所示的机器学习解决方案的执行流程进行说明。

（1）决策实体的确定

15 可以由计算图引擎构建图 5 所示的计算图（对应于上文述及的第一计算图）。

 需要注意的是，图 5 所示的计算图中的决策实体 1、决策实体 2 当前都处于待定状态；计算图中的目标实体 1 当前都处于待优化状态；计算图中的模型实体 1 当前都处于待训练状态；计算图中的数据实体 1、数据实体 3、数据实体 4 是原始数据实体。

20 一旦图 4 中的步骤 6 被执行，目标实体 1 执行其“优化”动作，这个动作将导致：

 1. 计算图引擎逆着虚线箭头方向找到所有原始数据实体（数据实体 1、3、4）、原始属性实体（图中未体现）、决策实体（决策实体 1、2）；

25 2. 对 1 中找到的所有原始数据实体（数据实体 1、3、4）、原始属性实体（图中未体现），执行它们的“获得实际数据”动作；

 3. 对 1 中找到的所有决策实体（决策实体 1、2），执行它们的“生成决策”动作（这一动作的具体行为由后端引擎中的优化引擎的算法决定）

30 4. 顺箭头顺序执行计算图，执行过程中的每一个函数（功能由其底层的机器学习引擎或自动机器学习引擎实现），执行过程中每一个生成数据实体和生成属性实体的“获得实际数据”动作（生成数据实体和生成属性实体的“获得实际数据”与原始实体不同，后者主要是根据描述文件获得对应的实际数据，而前者则需要进行计算图描述的计算，期间会利用到原始实体的实际数据），并对执行过程中经过的模型实体（模型实体 1）执行其“训练”动作。对每个模型，如果以其为起点逆虚线箭头方向，可以找到状态为“待定”的决策实体，则保留其状态为“待训练”；否则设置其状态为“已训练”

35 5. 对执行“优化”的目标实体本身，执行其“计算当前值”动作，优化引擎将获得这个值，以给出下一组候选决策。

 6. 重复执行 3~5 直到优化引擎或用户中止该流程，每次，决策实体的“生成决策”动作将根据优化引擎的算法产生不同的值。

上述流程结束后，与目标实体 1 的所有决策实体（决策实体 1、2）作出了决策。此时决策实体 1、2 状态变为“确定”。

(2) 模型训练过程

5 在执行图 4 中步骤 6 之后，所有与模型实体 1 相关（即与目标实体 1 相关的，因为目标实体 1 是基于模型实体 1 定义的）的决策实体都已变为“确定”状态。此时用户可以执行图 4 中示出的步骤 7，令模型实体 1 执行“训练函数”，并根据结果修改模型实体 1。

10 作为示例，可以构建出图 6 所示的计算图（对应于上文述及的第二计算图，不包含所有加粗部分，忽略了目标实体 1，因为它已与模型训练无关）。需注意图 6 中的决策实体 1、2 处于“确定”状态，但是模型实体 1 仍处于“待训练”状态。

一旦步骤 7 被执行，模型实体 1 执行其“训练”动作，这个动作将导致：

1. 以模型实体 1 为起点，逆虚线箭头找到所有原始数据实体（数据实体 1、3、4）、原始属性实体（图中未体现）、决策实体（决策实体 1、2）；
- 15 2. 对 1 中找到的所有原始数据实体（数据实体 1、3、4）、原始属性实体（图中未体现），执行它们的“获得实际数据”动作；

3. 对 1 中找到的所有决策实体（决策实体 1、2），执行它们的“获取最优决策”动作；

- 20 4. 顺箭头顺序执行计算图，执行过程中的每一个函数（功能由其底层的机器学习引擎或自动机器学习引擎实现，执行过程中每一个生成数据实体和生成属性实体的“获得实际数据”动作，执行过程中经过的模型实体（模型实体 1）执行其“训练”动作，并将其状态设置为“已训练”。

上述流程结束后，模型实体 1 处于“已训练”状态。完成图 4 所示的解决方案的开发。

25 需要说明的是，图 5 中通过数据划分函数划分得到的数据实体 5 和数据实体 6 是分别作为训练数据和验证数据的。图 6 中通过数据划分函数划分得到的数据实体 5 和数据实体 6 是均作为参与模型训练的训练数据的，也即可以认为数据实体 6 是一个空数据，通过数据划分函数，将所有数据都划分到了数据实体 5 中。在本公开的实施例中，训练数据根据机器学习模型的应用场景的不同而不同。例如，在图像处理的 OCR 场景中，所获取的数据记录是图像数据；而在金融科技领域的反洗钱、反欺诈相关的场景中，训练数据是银行用户的交易流水数据以及该用户本身相关的数据。本领域的技术人员能够理解不同场景下的训练数据的不同。

35 即本领域的技术人员可以理解，当机器学习模型应用于某个特定场景时，该机器学习模型是基于该场景对应的训练样本数据集训练而得的。例如，对于商品价格预测，对应的训练数据是该商品的历史数据（例如以商品历史上被售出时其自身的属性、季节、库存量等作为样本的特征，以被售出的价格作为标签），相应地，商品价格预测场景下，训练数据由商品当前的相关信息构成，基于训练数据构造训练样本，例如以商品当前的自身属性、季节和库

存量等作为预测样本的特征，将训练样本输入机器学习模型后得到模型输出的预测价格。其他场景类似，这里不再一一赘述。

5 自动机器学习的最大特点是根据实际任务（往往依赖于实际数据）来确定某些决策。在本公开中，为了开发一个自动机器学习解决方案，用户只需要将其认为重要的决策设定为决策实体，并且在发布、分享、部署与解决方案对应的本体模型（例如与图 4 所示的解决方案对应的本体模型）前保证这些决策实体处于“待定”状态。面对新的实际任务时，执行对应目标实体的“优化”动作，这样就可以令这些决策针对实际数据作出调整。

10 用户可以通过本公开涉及的交互式开发界面，以交互方式进行以图 4 为例的机器学习或自动机器学习解决方案的开发。开交互式界面支持所有本公开涉及的领域专业语言，并且可以提供以下功能。

1) 获取反馈功能：用户可以查看每一个实体的具体定义及其所有属性、查看数据实体和属性实体对应的实际数据、查看模型实体/决策实体/目标实体的状态、查看变量/常量实体的值等功能。可选地，交互式开发界面可以用图形形式展示本体模型（这里的本体模型可以是原始的本体模型，也可以是与整个解决方案对应的完整本体模型）、属性实体（数据实体的属性）。对属性实体（包含数据实体的属性）可以通过图形形式展示其产生过程。

2) 获取建议功能：本公开涉及的领域专门语言提供一系列自动机器学习函数，可以根据原始的本体模型（主要是数据实体及其关系）给出一个包含全部或部分可行操作的集合（此处可行操作指一种基于当前本体模型合法的、具体的操作，例如：以某些数据实体和变量/常量实体为输入，调用某个特定的特征工程函数，生成一个或一组属性主体）作为变量实体。用户可以查看、修改、采纳该变量实体中的全部或部分操作。由此，可以便于用户针对原始的本体模型设置机器学习解决方案。

25 本公开还可以将图 4 所示的本体模型导出为文件，或将文件导入为对应的本体模型。因为本体模型（例如图 4 所示）包含了完整的机器学习或自动机器学习方案，所以导出和导入本体模型就可以支持方案的分享与重用。

30 可选地，本体模型还可以包含版本管理功能，支持分支（branch）、检查不同（diff）、合并（merge）、查看历史（log）等操作。这些操作的含义可以参见 git。因为本体模型可以对应于一个计算图，所以理论上可以实现通过比较两个计算图之间的区别、合并两个计算图来实现版本管理。

35 本公开的机器学习建模过程的实现方法，还可以实现为一种机器学习建模过程的实现系统。图 7 示出了根据本公开示例性实施例的机器学习建模过程的实现系统的结构框图。其中，机器学习建模过程的实现系统的功能单元可以由实现本公开原理的硬件、软件或硬件和软件的结合来实现。本领域技术人员可以理解的是，图 7 所描述的功能单元可以组合起来或者划分成子单元，从而实现上述发明的原理。因此，本文的描述可以支持对本文描述的功能单元的任何可能的组合、或者划分、或者更进一步的限定。

下面就机器学习建模过程的实现系统可以具有的功能单元以及各功能单

元可以执行的操作做简要说明，对于其中涉及的细节部分可以参见上文相关描述，这里不再赘述。

参见图 7，机器学习建模过程的实现系统 700 包括交互装置 710 和机器学习处理装置 720。

- 5 交互装置 710 用于获取用户针对至少部分机器学习建模过程设定的配置信息，配置信息用于表征至少部分机器学习建模过程的执行策略，配置信息包括执行策略已确定的第一部分和执行策略待确定的第二部分。关于配置信息可以参见上文相关描述，此处不再赘述。

- 10 作为示例，交互装置 710 还用于获取数据描述信息，数据描述信息描述了一种或多种数据、数据的属性以及数据和/或属性的关系，以及基于数据描述信息，得到用于结构化地表征数据、属性以及关系的数据模型。关于数据描述信息、数据模型可以参见上文相关描述，此处不再赘述。

- 15 数据描述信息可以是用户基于第一规则语言编写的，交互装置 710 可以使用能够识别第一规则语言的读取函数读取数据描述信息，并基于数据描述信息得到数据模型。

可选地，交互装置 710 还用于向用户输出建议信息，建议信息用于表征针对数据模型的可行操作。

配置信息可以是用户根据数据模型设定的，配置信息用于表征基于数据模型所表征的数据和/或属性实现至少部分机器学习建模过程的执行策略。

- 20 机器学习处理装置 720 用于利用自动机器学习技术确定第二部分的执行策略，以及基于第一部分的执行策略和所确定的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。关于机器学习处理装置 720 确定第二部分的执行策略的具体过程，可以参见上文相关描述，此处不再赘述。

- 25 作为示例，机器学习处理装置 720 可以包括解释器和后端引擎，解释器用于将配置信息翻译为后端引擎对应的语言或指令，后端引擎用于基于语言或指令，确定第二部分的执行策略，以及基于第一部分的执行策略和所确定的第二部分的执行策略，执行至少部分机器学习建模过程，得到与至少部分机器学习建模过程对应的结果。

- 30 后端引擎可以包括计算图引擎和执行引擎，执行引擎可以包括优化引擎、机器学习引擎、自动机器学习引擎中的一项或多项。

可以由计算图引擎基于语言或指令生成第一计算图，第一计算图用于表征确定第二部分的执行策略的实现流程，由计算图引擎通过调用执行引擎执行第一计算图，确定第二部分的执行策略。

- 35 还可以由计算图引擎基于语言或指令生成第二计算图，第二计算图用于表征至少部分机器学习建模过程的实现流程，并由计算图引擎通过调用执行引擎执行第二计算图，得到与至少部分机器学习建模过程对应的结果。

应该理解，根据本公开示例性实施例的机器学习建模过程的实现系统 700 的具体实现方式可参照结合图 1 描述的相关具体实现方式来实现，在此不再

赘述。

以上参照图 1 到图 7 描述了根据本公开示例性实施例的机器学习建模过程的实现方法、机器学习建模过程的实现系统。应理解，上述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，在所述计算机可读介质上记录有用于执行图 1 示出的机器学习建模过程的实现方法的计算机程序。

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，所述计算机程序除了可用于执行除了图 1 示出的步骤之外，还可用于执行除了上述步骤以外的附加步骤或者在执行上述步骤时执行更为具体的处理，这些附加步骤和进一步处理的内容已经参照图 1 进行了描述，这里为了避免重复将不再进行赘述。

应注意，根据本公开示例性实施例的机器学习建模过程的实现系统，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，图 7 所示的各个装置也可以通过硬件、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。

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，执行机器学习建模过程的实现方法。

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。

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机器学习建模过程的实现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。

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。

5 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。

10 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机器学习建模过程的实现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。

15 例如，如上所述，根据本公开示例性实施例的机器学习建模过程的实现系统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，执行上文述及的机器学习建模过程的实现方法。

20 以上描述了本公开的各示例性实施例，应理解，上述描述仅是示例性的，并非穷尽性的，本公开不限于所披露的各示例性实施例。在不偏离本公开的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。因此，本公开的保护范围应该以权利要求的范围为准。

工业实用性

25 本公开提供的机器学习建模过程的实现方法和系统，通过参考用户的机器学习经验以及对业务的理解，从用户设定的配置信息中获取用户能够确定执行策略的第一部分；对于用户无法确定执行策略的第二部分，再利用自动机器学习技术确定该部分的执行策略；从而可以将人与机器的优势结合起来，提高机器学习过程的实现效率。