

预测方法、装置、可读介质及电子设备

技术领域

本发明涉及能源和物联网技术领域，尤其涉及一种设备故障预测方法和装置、设备状态
5 预测方法和装置、可读介质及电子设备。

背景技术

智能制造系统中的工业设备一旦发生故障，将会影响智能制造系统的正常工作，造成十分严重的影响，因此通常需要对工业设备进行设备故障预测。目前，在对某一工业设备进行
10 预测时，由于该工业设备的运行数据没有标签数据，因此需要采集若干个其他工业设备的历史运行数据以及历史运行数据对应的标签数据，采用机器学习方法建立这些历史运行数据与设备故障预测之间的映射关系，得到设备故障预测模型，使用设备故障预测模型实现该工业设备的设备故障预测。但是，上述技术方案为了学习到有效的模型，在对多个工业设备的设备故障预测时，可能需要共享工业设备的历史运行数据以及历史运行数据对应的标签数据，
15 因此导致工业设备之间的数据安全性较低。

设备预测性维护是根据设备运行状态的特征信息预测设备失效的概率或设备的剩余使用寿命等。其中基于数据驱动的方法是使用设备历史运行数据，包括设备各个传感器的量测，通过机器学习方法建立这些量测与设备发生故障概率之间的映射关系。为了学习到有效的模型，通常需要高质量的标注数据，也就是需要大量的故障标签数据；单台设备的故障数据有
20 限，需要联合多台同类型的设备获得足够的数据。已有的技术方案能够解决设备自身故障预测的问题，但是对不同设备之间数据共享，目前已经有的技术方案，是通过直接共享物联网中各个设备的数据，并未考虑物联网中各设备之间保护数据隐私的需求。

此外，进行设备状态的自动监测，已经成为保障设备正常运行的一种重要技术手段。设备状态预测是根据设备运行数据预测设备故障的概率或设备的剩余使用寿命。目前，对于难以获得设备状态标签的待预测设备，通过获取其他设备的历史运行数据以及历史运行数据对应的设备状态标签数据，从而获取大量的标注数据，之后基于机器学习方法建立这些历史运行数据与设备状态标签数据之间的映射关系，获得设备状态预测模型，基于设备状态预测模型，从而实现对待预测设备的设备状态预测。但是，不同设备所产生的历史运行数据的数据分布有所差异，导致利用上述技术方案所得到的设备状态预测模型的精度相对较低。
25 30

发明内容

本发明针对现有技术中存在的上述技术问题提供了一种设备故障预测方法、装置、计算机可读存储介质及电子设备，可将非共享数据迁移到目标设备上，建立目标设备的特征数据和设备故障之间的关系，无需共享的设备之间的特征数据，确保了数据安全。

35 第一方面，本发明提供了一种设备故障预测方法，包括：确定目标设备的特征信息和目标设备对应的检测点数据信息；基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型；根据所述特征数据概率分布模型和所述检测点数据的概率分布模型，确定所述非共享数据的权重；根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障
40 标签，建立联邦学习模型；根据所述联邦学习模型进行所述目标设备的设备故障预测。

第二方面，本发明提供了一种设备故障预测装置，包括：信息确定模块，用于确定目标设备的特征信息和目标设备对应的检测点数据信息；概率模型确定模块，用于基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型；权重确定模块，用于根据所述特征数据概率分布模型和所述检测点数据的概率分布模型，确定所述非共享数据的权重；模型建立模块，
45 用于根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签，

建立联邦学习模型；预测模块，用于根据所述联邦学习模型进行所述目标设备的设备故障预测。

第三方面，本发明实施例提供了一种设备故障预测方法，该设备故障预测方法包括：根据目标设备属性，获取用于对目标设备建立预测模型的训练数据集，其中所述数据集中的样本数据为共享数据；计算所述训练数据集中每条样本数据的权重；利用所述权重训练得到目标设备的故障预测局部模型；基于所述故障预测局部模型与联合学习算法，建立联合模型；根据所述联合模型对所述目标设备进行故障预测。

第四方面，本发明实施例提供了一种设备故障预测装置，该设备故障预测装置包括：数据获取模块、权重计算模块、局部模型训练模块、联合模型建立模块和故障预测模块，其中，所述数据获取模块，用于根据目标设备属性，获取用于对目标设备建立预测模型的训练数据集，其中所述数据集中的样本数据为共享数据；所述权重计算模块，用于计算所述训练数据集中每条样本数据的权重；所述局部模型训练模块，用于利用所述权重训练得到目标设备的故障预测局部模型；所述联合模型建立模块，用于基于所述故障预测局部模型与联合学习算法，建立联合模型；所述故障预测模块，用于根据所述联合模型对所述目标设备进行故障预测。

第五方面，本发明提供了一种设备状态预测方法，包括：获取目标设备对应的至少两个特征数据、至少两个参考设备分别对应的至少两个模型训练数据以及各个所述模型训练数据分别对应的设备状态标签，所述目标设备和所述参考设备的设备种类相同；根据各个所述特征数据，确定所述模型训练数据对应的误差权重；根据各个所述模型训练数据、各个所述模型训练数据分别对应的设备状态标签以及各个所述模型训练数据分别对应的误差权重进行模型训练，以确定设备状态预测模型，所述设备状态预测模型用于所述目标设备的设备状态预测。

第六方面，本发明提供了一种设备状态预测装置，包括：获取模块，用于获取目标设备对应的至少两个特征数据、至少两个参考设备分别对应的至少两个模型训练数据以及各个所述模型训练数据分别对应的设备状态标签，所述目标设备和所述参考设备的设备种类相同；权重确定模块，用于根据各个所述特征数据，确定所述模型训练数据对应的误差权重；训练模块，用于根据各个所述模型训练数据、各个所述模型训练数据分别对应的设备状态标签以及各个所述模型训练数据分别对应的误差权重进行模型训练，以确定设备状态预测模型，所述设备状态预测模型用于所述目标设备的设备状态预测。

第七方面，本发明提供了一种计算机可读存储介质，包括执行指令，当电子设备的处理器执行所述执行指令时，所述处理器执行如第一方面中任一所述的方法或如第二方面中任一所述的方法。

第八方面，本发明提供了一种电子设备，包括处理器以及存储有执行指令的存储器，当所述处理器执行所述存储器存储的所述执行指令时，所述处理器执行如第一方面中任一所述的方法或如第二方面中任一所述的方法。

本发明提供了一种设备故障预测方法和装置、设备状态预测方法和装置、计算机可读存储介质及电子设备，该方法通过确定目标设备的特征信息和目标设备对应的检测点数据信息，然后，基于目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型，然后根据特征数据概率分布模型和检测点数据的概率分布模型，确定非共享数据的权重，然后，根据非共享数据、非共享数据的权重和非共享数据对应的设备故障标签，建立联邦学习模型，最后，根据联邦学习模型进行目标设备的设备故障预测。本发明提供的技术方案，通过将非共享数据迁移到目标设备上，从而建立目标设备的特征数据和设备故障之间的关系，无需共享设备之间的特征数据，确保了数据安全。

上述的非惯用的优选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

附图说明

为了更清楚地说明本发明实施例或现有的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明中记载的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

- 5 图 1 为本发明一实施例提供的一种设备故障预测方法的流程示意图；
图 2 为本发明一实施例提供的另一种设备故障预测方法的流程示意图；
图 3 为本发明一实施例提供的一种设备故障预测方法的流程示意图；
图 4 为本发明一实施例提供的另一种设备故障预测方法流程示意图；
10 图 5 为本发明一实施例提供的一种设备状态预测方法的流程示意图；
图 6 为本发明一实施例提供的另一种设备状态预测方法的流程示意图；
图 7 为本发明一实施例提供的一种设备故障预测装置的结构示意图；
图 8 为本发明一实施例提供的一种设备故障预测装置的结构示意图；
图 9 为本发明一实施例提供的一种设备状态预测装置的结构示意图；
15 图 10 为本发明一实施例提供的一种电子设备的结构示意图。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合具体实施例及相应的附图对本发明的技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本发明一部分实施例，
20 而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

图 1 为本发明一实施例提供的一种设备故障预测方法的流程示意图。如图 1 所示，本发明实施例提供了一种设备故障预测方法，包括如下多个步骤：

步骤 101，确定目标设备的特征信息和目标设备对应的检测点数据信息。

- 25 具体地，目标设备为智能制造系统的任意一种工业设备，本发明实施例对此不作具体限定，优选地，目标设备为能源设备，比如，燃气锅炉、内燃机、汽轮机、热电联产设备、光伏设备等。

具体地，目标设备的特征信息指的是目标设备的特征数据的属性或功能。

具体地，检测点数据信息指的是检测点数据的属性或功能。

- 30 需要说明的是，检测点数据包括非共享数据。

步骤 102，基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型。

- 具体地，目标设备的特征数据有多个，每个特征数据包括若干个特征分别对应目标设备的特征值，其中，若干个特征为设备故障的影响因素，具体需要结合实际场景确定，比如，
35 目标设备为燃气锅炉，则若干个特征包括但不限于燃气流量、燃气温度、排烟温度、烟气流量、燃气压力、开停机状态、烟气湿度以及烟气压力等。

具体地，特征数据服从连续概率分布，因此，特征数据概率分布模型可以是正态分布模型，也可以是指数分布模型，本发明实施例对此不做限定，优选地，特征数据概率分布模型是由多个正态分布模型混合成的混合高斯模型，其中，正态分布模型又称高斯分布模型。

- 40 具体地，检测点数据的概率分布模型与特征数据概率分布模型相同，优选地，可以是混合高斯模型，当然，也可以是其他概率分布模型，具体需要结合实际情况确定，此处不做具体限定。

- 本发明一实施例中，所述基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型
45 确定所述特征数据概率分布模型，包括：基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据和具有非共享数据的检测点数据；根据特征数据参

数模型计算所述特征数据的数据分布，并将确定好参数的特征数据参数模型确定为所述特征数据概率分布模型；根据检测点数据参数模型计算所述检测点数据的数据分布，并将确定好参数的检测点数据参数模型确定为所述检测点数据的概率分布模型。

具体地，特征数据参数模型可以是高斯模型也可以是混合高斯模型，优选混合高斯模型。

- 5 检测点数据参数模型与特征数据参数模型相同，本发明实施例对此不做过多赘述。

具体地，特征数据概率分布模型是通过将各个特征数据代入到混合高斯模型中进行计算来确定模型参数，其中计算方法可以是 EM 算法，将确定好模型参数的混合高斯模型确定为特征数据概率分布模型。

- 10 检测点数据的概率分布模型，是通过将多个检测点数据代入到混合高斯模型中进行计算来确定模型参数，其中计算方法可以是 EM 算法，将确定好模型参数的混合高斯模型确定为检测点数据的概率分布模型。

步骤 103，根据所述特征数据概率分布模型和所述检测点数据的概率分布模型，确定所述非共享数据的权重。

- 15 该实施例中，通过非共享数据的权重，从而将检测点的非共享数据迁移到目标设备上，换言之，建立了检测点和目标设备之间的数据关联，检测点不会直接获取到目标设备的特征数据，即目标设备和检测点之间不存在数据共享，从而确保了数据安全。

- 20 本发明一个实施例中，所述根据所述特征数据概率分布模型和所述检测点数据的概率分布模型，确定所述非共享数据的权重，包括：根据所述特征数据概率分布模型，确定所述非共享数据的目标设备分布概率；根据所述检测点数据的概率分布模型，确定所述非共享数据的检测点分布概率；将所述非共享数据的目标设备分布概率和所述非共享数据的检测点分布概率的比值，确定为所述非共享数据的权重。

该实施例中，通过将非共享数据的目标设备分布概率和非共享数据的检测点分布概率的比值，确定非共享数据的权重，从而将检测点上的非共享数据迁移到目标设备上，同时无需共享目标设备的特征数据，确保了数据安全。

- 25 具体地，将非共享数据代入特征数据概率分布模型中，特征数据概率分布模型输出的值即为非共享数据的目标设备分布概率。

具体地，将非共享数据代入检测点数据的概率分布模型，检测点数据的概率分布模型输出的值即为共享数据的检测点分布概率。

- 30 步骤 104，根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签，建立联邦学习模型。

该实施例中，通过非共享数据、非共享数据的权重和非共享数据对应的设备故障标签，建立联邦学习模型，实现将非共享数据到目标设备的迁移，无需共享目标设备的特征数据，从而确保了数据安全。

- 35 需要说明的是，目标设备的特征数据和检测点数据分布在物联网中的不同的检测点，共享数据进行模型训练会产生数据安全问题，本发明实施例通过检测点数据中的非共享数据、非共享数据的权重以及非共享数据对应的设备故障标签，实现将非共享数据迁移到目标设备上，使得检测点之间不存在数据共享，避免了直接共享数据带来的数据安全问题。其中，检测点是能进行数据处理以及数据交互的节点，包括但不限于边缘服务器、边缘网关以及边缘控制器中的任意一种或多种。

- 40 具体地，设备故障标签可以是设备运行状态的特征信息预测设备失效的概率，比如，设备故障标签可以是故障类型，还可以是故障程度，本发明实施例对此不作限定，具体需要结合实际需求确定。需要说明的是，本发明实施例并不意图对每个非共享数据分别对应的设备故障标签的获取方法进行限定，也可以是人工标注，也可以是规则标注，还可以是聚类标注，上述标注方法均为现有技术，此处不做过多赘述。值得注意的是，本发明实施例提供的设备故障预测方法，也可以用来预测设备剩余使用寿命，具体地，通过设备剩余使用寿命替换设备故障标签即可。
- 45

在一个实施例中,所述根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签,建立联邦学习模型,包括:根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签,确定检测点局部模型;根据至少两个所述检测点局部模型,建立联邦学习模型。

5 具体地,基于非共享数据、非共享数据对应的权重和设备故障标签进行模型训练,确定检测点局部模型,然后,对多个检测点局部模型进行融合,建立联邦学习模型。

具体地,非共享数据以及非共享数据的权重能够实现将检测点的非共享数据迁移到目标设备上,后续,通过多个非共享数据各自的权重调整模型参数,从而使得调整后的模型参数能够反映出目标设备的特征数据和设备故障之间的关系,不涉及到检测点数据和目标设备的特征数据的共享,从而确保了数据安全。

10 本发明的一个实施例中,所述根据至少两个所述检测点局部模型,建立联邦学习模型,包括:基于联合学习算法,对至少两个所述检测点局部模型进行反复迭代,并基于至少两个所述检测点的迭代好的检测点局部模型,建立联邦学习模型。

具体地,联合学习算法指的是将各个检测点局部模型中的检测点局部模型参数发送到目标设备的检测点,目标设备的检测点对各个检测点局部模型参数进行平均或加权平均得到联邦学习模型参数,基于联邦学习参数进行检测点局部模型更新迭代。

具体地,通过将各个检测点的检测点局部模型的模型参数发送到目标设备的检测点,目标设备的检测点将来自各检测点的检测点局部模型的模型参数进行融合得到联邦学习模型参数,并将融合后的联邦学习模型参数再分发给各个检测点的检测点局部模型,之后根据各个检测点的非共享数据和融合后的联合模型参数进行本地训练得到更新后的模型参数,将更新后的模型参数发送给目标设备的检测点,按照上述方法反复迭代,直到迭代次数满足预设次数,或者,迭代后的检测点局部模型的模型误差满足预设值。将迭代好的检测点局部模型参数进行融合以得到联邦学习模型参数,进而得到联邦学习模型。其中,各个检测点局部模型的模型参数的融合的方式可以是平均也可以是加权平均,本发明实施例对此不做限定。

20 具体地,检测点局部模型可以是神经网络模型,也可以是回归模型,具体需要结合实际需求确定。

具体地,可通过如下方式对检测点局部模型进行迭代:

A1、根据多个非共享数据、每个非共享数据分别对应的设备故障标签以及每个非共享数据分别对应的权重进行模型训练,以确定检测点局部模型;

30 A2、判断检测点局部模型的模型误差是否满足迭代条件,若是,则将检测点局部模型确定为最终模型发送给目标设备的检测点,若否,则执行 A3;

A3、将检测点局部模型的模型参数发送给目标设备的检测点;

A4、接收目标设备的检测点发送的融合后的模型参数,并根据多个非共享数据、多个非共享数据分别对应的设备故障标签以及多个非共享数据分别对应的权重,对融合后的模型参数进行调整,以确定调整后的模型参数,并将调整后的模型参数替换检测点局部模型的模型参数,执行 A2。

步骤 S105,根据至少两个所述检测点的局部模型,建立联邦学习模型。

具体地,采集目标设备的当前特征数据,将当前特征数据代入联邦学习模型中,即可判断目标设备是否故障。

40 为了更好的理解目标设备的检测点和非共享数据的检测点之间的数据处理过程,举例来说,假设 3 个非共享数据的检测点设为 A、B 以及 C,目标设备的节点设为 D,以 A 为例进行说明,A 根据混合高斯模型确定好非共享数据的概率分布模型,并接收 D 发送来的特征数据概率分布模型,从而确定非共享数据对应的权重,B 以及 C 按照上述 A 相似的处理过程,得到 B 及 C 中的每个非共享数据分别对应的权重,之后,A、B、C 分别根据各个非共享数据、非共享数据对应的设备故障标签以及每个非共享数据分别对应的权重进行模型训练,以得到检测点局部模型,之后,D 对 A、B、C 发来的检测点局部模型进行联邦学习,即可得

到目标设备的联邦学习模型。

通过以上技术方案可知，本发明实施例至少存在以下有效效果：通过特征数据概率分布模型和非共享数据的概率分布模型确定非共享数据的权重，以得到目标设备的检测点局部模型，建立非共享数据的检测点和目标设备之间的数据关联，将非共享数据迁移到目标设备上，
5 同时，基于非共享数据建立联邦学习模型，利用该联邦学习模型实现目标设备的设备故障预测，非目标设备和目标设备之间无需共享设备的特征数据，从而确保数据安全。

图 1 所示仅为本发明所述方法的基础实施例，在其基础上进行一定的优化和拓展，还能够得到所述方法的其他优选实施例。

为了更加清楚的说明本发明的技术方案，请参考图 2。图 2 为本发明一实施例提供的另一种设备故障预测方法的流程示意图。本实施例在前述实施例的基础上，结合具体应用场景进行进一步的叙述。如图 2 所示，该设备故障预测方法具体可以包括如下各个步骤：

步骤 201，确定目标设备的特征信息和目标设备对应的检测点数据信息。

具体地，假设目标设备对应在物联网中的检测点为 n ，有 N 个非共享数据的检测点，分别对应在物联网中的检测点分别为 $n1$ 、 $n2$ 、 $\dots$ 、 nN ， $n1$ 、 $n2$ 、 $\dots$ 、 nN 分别能够获取到非共享数据， n 能够获取到目标设备所有的特征数据。需要说明的是， $n1$ 、 $n2$ 、 $\dots$ 、 nN 得到检测点局部模型的过程相似，下述仅以 $n1$ 的处理过程进行描述。
15

步骤 202，基于所述目标设备的特征信息和混合高斯模型，计算所述特征信息对应的特征数据的数据分布，并将确定好模型参数的混合高斯模型确定为所述特征数据概率分布模型。

n 根据高斯混合模型，将各个特征数据代入混合高斯模型中对各个特征数据计算，从而得到各个特征数据概率分布模型，并将各个特征数据概率分布模型分别送给 $n1$ 、 $n2$ 、 $\dots$ 、 nN 。
20

$n1$ 、 $n2$ 、 $\dots$ 、 nN 得到非共享数据的概率分布模型的过程相似，下述以 $n1$ 的处理过程进行描述。

步骤 203，基于目标设备对应的检测点数据信息和混合高斯模型计算所述检测点数据信息对应的具有非共享数据的检测点数据的数据分布，并将确定好模型参数的混合高斯模型确定为所述非共享数据的概率分布模型。
25

$n1$ 获取多个非共享数据，将非共享数据代入混合高斯模型进行计算，得到 $n1$ 的非共享数据的概率分布模型。

步骤 204，根据所述特征数据概率分布模型，确定所述非共享数据的目标设备分布概率，根据所述检测点数据的概率分布模型，确定所述非共享数据的检测点分布概率。
30

$n1$ 接收到特征数据对应的概率分布模型，将其获得的非共享数据代入特征数据概率分布模型中进行计算，得到非共享数据的目标设备分布概率。

$n1$ 将非共享数据代入到非共享数据的概率分布模型中进行计算，得到非共享数据的检测点分布概率。
35

步骤 205，将所述非共享数据的目标设备分布概率和所述非共享数据的检测点分布概率的比值，确定为所述非共享数据的权重。

$n1$ 通过将非共享数据相对于目标设备的分布概率和相对于非共享数据的检测点的分布概率的概率比值确定为非共享数据的权重，基于相同的方法，得到每个非共享数据的权重，这样就可以实现建立起目标设备和非共享数据的检测点之间的数据联系。

步骤 206，根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签，确定所述检测点局部模型。
40

$n1$ 根据非共享数据、非共享数据的权重及非共享数据对应的设备故障标签进行模型训练，确定目标设备的检测点局部模型。 $n2$ 、 $\dots$ 、 nN 按照 $n1$ 相似的方法分别确定出目标设备的检测点局部模型。

步骤 207，基于联合学习算法，对至少两个所述检测点的检测点局部模型进行反复迭代，并基于至少两个所述检测点的迭代好的检测点局部模型，建立联邦学习模型。
45

n 与 n1、n2、……、nN，采用联合学习算法，对 n1、n2、……、nN 分别训练得到的检测点局部模型进行模型迭代后发给 n，n 对 n1、n2、……、nN 分别训练得到的迭代好的检测点局部模型进行模型平均，建立目标设备的联邦学习模型。

步骤 208，将所述联邦学习模型进行所述目标设备的设备故障预测。

- 5 n 根据联邦学习模型进行目标设备的设备故障预测。

通过以上技术方案可知，本实施例存在的有益效果是：通过非共享数据的概率分布模型与特征数据概率分布模型的比值，得到非共享数据的权重，建立非共享数据的检测点和目标设备之间的数据关联，将非共享数据迁移到目标设备上，通过目标设备和非共享数据在物联网中的检测点实现目标设备的联邦学习模型的建立，从而实现目标设备和非共享数据的检测点之间不涉及到特征数据的共享，确保了数据安全。

图 3 为本发明一实施例提供的一种设备故障预测方法的流程示意图。如图 3 所示，该设备故障预测方法包括：

步骤 301，根据目标设备属性，获取用于对目标设备建立预测模型的训练数据集，其中所述数据集中的样本数据为共享数据；

- 15 步骤 302，计算所述训练数据集中每条样本数据的权重；

步骤 303，利用所述权重训练得到目标设备的故障预测局部模型；

步骤 304，基于所述故障预测局部模型与联合学习算法，建立联合模型；

步骤 305，根据所述联合模型对所述目标设备进行故障预测。

在该实施例中，训练数据集中的样本数据包含目标设备的特征数据和样本设备的特征数据、样本设备的故障数据；样本设备为与目标设备相关或相似的设备。故障预测局部模型和联合模型均为设备的特征数据与设备发生故障的关系。不同设备数据分布在物联网不同的节点，共享数据训练模型会产生数据安全问题，该实施例采用一种基于联合学习方式的样本迁移方法，用于设备预测性维护，可以在多台设备上采集到的数据联合学习，并迁移到目标设备上，用来训练针对目标设备的预测性维护模型，实现多方联合学习，保证各方数据不出本地，避免了直接共享数据带来的数据安全问题。

在本发明一个实施例中，所述计算所述训练数据集中每条样本数据的权重，包括针对每台所述样本设备：将所述样本设备的特征数据与所述目标设备的特征数据进行区分；根据区分后的特征数据，训练分类模型；根据训练后的分类模型，计算所述样本设备每条特征数据的权重。

30 由于样本设备和待预测设备可以分别分布物联网中的任何节点，为保证数据隐私，样本设备和待预测设备的数据不能共享。因此可以采用基于联合学习的分类模型，如在该实施例中使用基于联合学习的 XGBoost 模型。

在本发明一个实施例中，所述计算所述训练数据集中每条样本数据的权重，包括针对每台所述样本设备：将所述样本设备的特征数据标记为第一数据，将所述待预测设备的特征数据标记为第二数据；根据所述第一数据和所述第二数据，训练分类模型，所述分类模型为基于联合学习的分类模型；根据训练后的分类模型，计算所述样本设备每条特征数据的权重，所述权重的计算公式为：

$$\omega_i = \frac{P_{2i}}{P_{1i}}$$

其中， ω_i 为所述第一数据中第 i 条数据的权重， P_{1i} 为所述第 i 条数据属于所述样本设备的概率， P_{2i} 为所述第 i 条数据属于所述目标设备的概率。

40 在本发明一个实施例中，所述利用所述权重训练得到目标设备的故障预测局部模型，包括：根据每台所述样本设备的特征数据、所述样本设备每条特征数据的权重和每台所述样本设备的故障数据，使用神经网络分别在所述样本设备有权重的所述训练数据集上训练得到所述目标设备的故障预测局部模型。

在本发明一个实施例中，所述基于所述故障预测局部模型与联合学习算法，建立联合模型，包括：根据所述故障预测局部模型，使用联合学习算法反复迭代，获得所述样本设备在所述训练数据集上关于所述目标设备的联合模型。

在本发明一个实施例中，所述目标设备和各个所述样本设备均为物联网中的边缘节点，所述目标设备的特征数据不对其他所述样本设备暴露，每个所述样本设备的特征数据和故障数据不对其他所述样本设备和所述目标设备暴露。

图4为本发明一实施例提供的另一种设备故障预测方法流程示意图。如图4所示，该设备故障预测方法包括：

步骤401，获取每台样本设备的传感器特征数据，以及相应的故障数据；采集目标设备的传感器特征数据；

步骤402，分别为目标设备和样本设备的特征数据打标签，得到标签数据；

步骤403，基于步骤402的标签数据，采用联合学习的方式训练分类器（二分类）；

步骤404，根据步骤403的分类器，计算样本设备特征数据的权重；

步骤405，重复以上步骤计算各样本设备特征数据的权重；

步骤406，基于各个样本设备的数据，采用联合学习的方式建立传感器特征数据和设备发生故障之间的关系；

步骤407，将步骤406的模型用于目标设备发生故障概率的预测。

假定具有设备A、B、C等各个测点的特征数据，以及相应的故障标注，也就是具有设备A、B、C的所有特征数据和故障数据（标签数据）；同时也具有设备D相同测点的特征数据，也就是具有设备D的所有特征数据，不具备故障数据，需要根据这些特征数据预测设备D的故障发生概率。

目标是使用设备A、B、C的特征数据、故障数据和设备D的特征数据学习到一个预测设备D发生故障的概率模型，用于设备D的故障预测。

需要注意的是在此过程中，假定A、B、C、D等均为物联网中的边缘节点，各方都有数据隐私和安全的需求，在训练过程中自身特征数据和故障数据均不能对外暴露。

在该实施例中，设备故障预测方法包括：

1. 收集训练数据集训练数据集包括设备A、B、C的所有特征数据以及故障数据，还有设备D的特征数据；

2. 计算设备A、B、C每条特征数据的权重，以下以设备A为例，B、C情况相同，计算步骤如下：

a) 为设备A和设备D的特征数据打标签，假定设备A的标签为0，设备D的标签为1；

b) 基于上述设备A和设备D的标签数据，训练分类器（二分类）（本例中使用基于联合学习的XGBoost模型，实际中不限使用XGBoost，可以使用任何基于联合学习的概率分类模型）；

c) 对于设备A中的每条数据x，根据上述训练的二分类分类器，计算数据x属于设备A的概率 $PA(x)$ ，以及属于设备D概率 $PD(x)$ ；计算数据x权重， $\omega(x)=PD(x)/PA(x)$ ；

3. 训练模型

a) 由以上步骤得到设备A、B、C的特征数据和故障数据，以及每条特征数据对应的权重；

b) 训练局部模型，使用神经网络或相关回归算法分别在上述设备A、B、C的有权重的数据集上训练得到设备D的故障预测局部模型（具体训练过程为标准过程，不再赘述）；

c) 计算联合模型，根据上述在设备A、B、C数据集上训练得到的设备D的故障预测局部模型；使用模型平均（联合学习算法）反复迭代，计算设备A、B、C数据集上关于设备D故障预测的联合模型也就是故障预测模型；在该实施例中，反复迭代的次数由模型的预设精度和预设迭代次数决定。

4. 将得到的联合模型用于设备D的故障预测。

本实施例中暂假定设备 A、B、C 三台设备参与联合训练，在实际中，联合设备的数量没有限制。

值得注意的是，本发明中提到的设备故障数据，同样可以替换为设备剩余使用寿命，从而对设备的剩余使用寿命进行预测。

5 图 5 为本发明一实施例提供的一种设备状态预测方法的流程示意图。如图 5 所述，该设备状态预测方法包括如下各个步骤：

步骤 501，获取目标设备对应的至少两个特征数据、至少两个参考设备分别对应的至少两个模型训练数据以及各个所述模型训练数据分别对应的设备状态标签，所述目标设备和所述参考设备的设备种类相同。

10 具体地，可以根据接入网络的参考设备上安装的传感器，获取参考设备的运行数据，根据参考设备的运行数据，从而确定出参考设备对应的多个模型训练数据。需要说明的是，基于传感器采集的数据获取模型训练数据为现有技术，本实施例对此不作过多赘述。

具体地，可以根据接入网络的目标设备安装的传感器，获取目标设备的运行数据，根据目标设备的运行数据，从而确定出目标设备对应的多个特征数据。需要说明的是，基于传感器采集的数据获取特征数据为现有技术，本实施例对此不作过多赘述。

15 具体地，模型训练数据包括若干个特征以及每个特征分别对应参考设备的一个或多个特征值，优选一个，其中，若干个特征为对设备状态具有影响的影响因素。模型训练数据和特征数据所对应的若干个特征相同，特征对应的特征值可能不同。这里，模型训练数据分别对应一个设备状态标签。需要说明的是，本发明实施例并不意图对参考设备的数量进行限定，具体需要结合实际场景确定。

20 在一个实施例中，设备状态标签包括但不限于设备故障信息或设备的剩余使用寿命。需要说明的是，本发明实施例并不意图对每个模型训练数据分别对应的设备状态标签的获取方法进行限定，也可以是人工标注，也可以是规则标注，还可以是聚类标注，上述标注方法均为现有技术，此处不做过多赘述。

25 需要说明的是，参考设备和目标设备的设备种类相同，从而确保能够对参考设备和目标设备进行数据关联的价值，进而确保后续确定出的设备状态预测模型的参考价值。比如，参考设备和目标设备均是燃气锅炉。

还需要说明的是，目标设备的特征数据没有设备状态标签，无法利用自身的数据完成模型训练，因此需要借助参考设备的模型训练数据，得到其对应的设备状态预测模型。

30 步骤 502，根据各个所述特征数据，确定所述模型训练数据对应的误差权重。

在一个实施例中，具体可通过如下方法确定模型训练数据对应的误差权重：

从各个特征数据中确定出模型训练数据对应的替换数据；确定模型训练数据在其所属的参考设备对应的各个模型训练数据中的第一分布信息以及替换数据在各个特征数据中的第二分布信息；根据第一分布信息和第二分布信息，确定模型训练数据对应的权重误差。

35 为了能够确定出参考设备的模型训练数据对应的误差权重，了解参数设备和目标设备之间的数据关系，需要从目标设备的多个特征数据中找到模型训练数据对应的替换数据，这里，替换数据优选多个特征数据中与模型训练数据相同的数据，若多个特征数据中不存在和模型训练数据相同的数据，此时，替换数据可以为多个特征数据中和模型训练数据最相似的数据，作为一种可能的情况，当多个特征数据中存在 2 个或 2 个以上可替换模型训练数据的特征数据时，此时，可以随机选择一个可替换模型训练数据的特征数据作为替换数据。

40 然后，确定模型训练数据在其所属的参考设备对应的各个模型训练数据中的第一分布信息。这里，第一分布信息指示了模型训练数据在其所属的参考设备对应的所有模型训练数据的分布情况，第一分布信息可以是分布概率，也可以是出现频率，此处不做具体限定。

45 然后，确定替换数据在各个特征数据中的第二分布信息。这里，第二分布信息指示了替换数据在目标设备的所有特征数据的分布情况，第二分布信息可以是分布概率，也可以是出现频率，此处不做具体限定。需要说明的是，第一分布信息和第二分布信息所包括的内容应

当统一，比如，第一分布信息和第二分布信息可以都包括分布概率，也可以都包括出现频率，此处不做具体限定。

需要说明的是，本实施例并不意图对确定分布信息的方法进行限定，任何现有技术中用于确定数据分布的方法即可，比如，可以是参数估计方法，也可以是非参数估计方法，还可以是半参数估计方法。

在一个实施例中，第一分布信息为第一分布概率，第二分布信息为第二分布概率，具体可通过如下方式确定第一分布概率及第二分布概率：

基于模型训练数据在其所属的参考设备对应的第一非参数估计方法，确定模型训练数据在其所属的参考设备对应的各个模型训练数据中的第一分布概率；基于第二非参数估计方法，确定特征数据在各个特征数据中的第二分布概率。

非参数估计方法不受总体分布的限制，不假定总体分布的具体形式，尽量从数据或样本本身获得所需要的信息，通过估计而获得分布的结构，对于减少偏差、提高预测精度以及了解数据的动态结构具有极其重要的作用，尤其是数据分布很不规则时，非参数估计方法相对于参数估计来说，更为精准，因此本实施例中采用非参数估计方法确定分布概率。

考虑到参考设备和目标设备分别对应的数据分布可能存在差异，因此，为了确保得到的分布概率的准确性，通过第一非参数估计方法确定参考设备对应的各个模型训练数据的概率密度函数，基于概率密度分布函数从而确定出参考设备对应的每个模型训练数据的第一分布概率。通过第二非参数估计方法确定目标设备对应的各个特征数据的概率密度函数，基于概率密度分布函数从而确定出目标设备对应的每个特征数据的第二分布概率。

考虑到设备的数据量和数据分布会存在差异，若采用相同的非参数估计方法，则可能会降低估算出的数据分布的准确性，因此，需要结合每个参考设备以及目标设备的数据量以及数据分布，选择合适的非参数估计方法进行数据分布的估算。当每个参考设备分别对应的各个模型训练数据的数据量，和目标设备对应的各个特征数据的数据量的差异较小时，各个参考设备分别对应的第一非参数估计方法和目标设备对应的第二非参数估计方法相同，作为一种可能的情况，第一非参数估计方法和第二非参数估计方法均为核密度估计方法。

在一个实施例中，当第一分布信息为第一分布概率，第二分布信息为第二分布概率时，具体可通过如下方式确定模型训练数据对应的误差权重：

针对每个模型训练数据，确定模型训练数据对应的替换数据的第二分布概率和模型训练数据的第一分布概率的分布概率比值，将分布概率比值确定为确定模型训练数据对应的误差权重。

在一个实施例中，当第一分布信息为第一出现频率，第二分布信息为第二出现频率时，具体可通过如下方式确定模型训练数据对应的误差权重：

针对每个模型训练数据，确定模型训练数据对应的替换数据的第二出现频率和模型训练数据的第一出现频率的出现频率比值，将出现频率比值确定为确定模型训练数据对应的误差权重。

步骤 503，根据各个所述模型训练数据、各个所述模型训练数据分别对应的设备状态标签以及各个所述模型训练数据分别对应的误差权重进行模型训练，以确定设备状态预测模型，所述设备状态预测模型用于所述目标设备的设备状态预测。

在一个实施例中，设备状态预测模型可以是神经网络模型，也可以是回归模型，此处不做限定，具体需要结合实际情况确定。

在一个实施例中，本实施例提供的设备状态预测方法应用于目标设备对应的边缘服务器或本地服务器，从而降低云端服务器的计算压力。

具体地，设备状态预测模型以多个模型训练数据以及每个模型训练数据分别对应的设备状态标签数据作为训练数据，通过每个模型训练数据分别对应的误差权重调整设备状态预测模型的模型参数，从而使得调整后的模型参数能够反映出目标设备的特征数据和设备状态之间的联系，确保设备状态预测模型的模型精度。

在一个实施例中，通过可通过如下实现方式确定设备状态预测模型：

A1、根据将模型训练数据代入待预测模型中的预测结果以及模型训练数据对应的设备状态标签，确定模型训练数据分别对应的第一误差，根据模型训练数据对应的第一误差以及误差权重，以确定模型训练数据对应的第二误差；

5 A2、判断是否满足迭代次数或者各个模型训练数据分别对应的第二误差是否满足预设条件，如果是，则将待预测模型确定为设备状态预测模型，如果否，则执行 A3；

A3、根据各个模型训练数据分别对应的第二误差，对待预测模型中的模型参数进行调整，以确定调整后的模型参数，并将待预测模型中的模型参数替换为调整后的模型参数，执行 A1。

10 这里，各个模型训练数据分别对应的第二误差是否满足预设条件可以是各个模型训练数据分别对应的第二误差的均值是否小于阈值，也可以是各个模型训练数据分别对应的第二误差的标准方差是否小于阈值，此处不做具体限定。

具体地，预测结果可以理解为设备状态信息，第一误差为预测结果和设备状态标签之间的接近程度，举例来说，预测结果可以是故障或者不故障，当预测结果是故障，设备状态标签也是故障时，第一误差可以是 0，否则为 1。

15 通过以上技术方案可知，本实施例存在的有益效果是：通过确定多个参考设备分别对应的多个模型训练数据分别相对于目标设备的误差权重，从而将多个参考设备分别对应的多个模型训练数据迁移到目标设备上，后续，利用多个参考设备分别对应的多个模型训练数据、每个模型训练数据对应的设备状态标签以及每个模型训练数据对应的权重误差进行模型训练，得到用于预测目标设备的设备状态的设备状态预测模型，得到的设备状态预测模型由于综合考虑了模型训练数据对于目标设备的误差权重，从而具有相对较高的精度，后续基于设备状态预测模型进行目标设备的设备状态预测时，能够得到较为准确的目标设备的设备状态。

图 5 所示仅为本发明所述方法的基础实施例，在其基础上进行一定的优化和拓展，还能够得到所述方法的其他优选实施例。

25 图 6 为本发明一实施例提供的另一种设备状态预测方法的流程示意图。本实施例在前述实施例的基础上，结合应用场景进行了更加具体的描述。如图 6 所示，该设备状态预测方法具体包括以下步骤：

步骤 601，获取目标设备对应的至少两个特征数据、至少两个参考设备分别对应的至少两个模型训练数据以及各个所述模型训练数据分别对应的设备状态标签，所述目标设备和所述参考设备的设备种类相同。

30 假设 3 个参考设备分别为锅炉 A、锅炉 B 以及锅炉 C，目标设备为锅炉 D，锅炉 A、锅炉 B 以及锅炉 C 分别对应有 N 个模型训练数据，每个模型训练数据分别对应有一个故障标签，故障标签可以是故障或者正常，锅炉 D 对应有 n 个特征数据，特征数据和模型训练数据分别对应的若干个特征相同，但是每个特征分别对应的特征值可能不同。

35 步骤 602，基于所述模型训练数据在其所属的参考设备对应的第一非参数估计方法，确定所述模型训练数据在其所属的参考设备对应的各个所述模型训练数据中的第一分布概率。

40 假设锅炉 A、锅炉 B 以及锅炉 C 分别对应的第一非参数估计方法均为核密度估计方法，针对锅炉 A，基于核密度估计方法，确定锅炉 A 对应有 N 个模型训练数据的概率密度分布函数，针对锅炉 A 中的每个模型训练数据，基于锅炉 A 对应有 N 个模型训练数据的概率密度分布函数，确定该模型训练数据在锅炉 A 中的所有模型训练数据中的第一分布概率 P1。锅炉 B 及锅炉 C 分别对应的每个模型训练数据的第一分布概率 P1 的计算方法和锅炉 A 相似，这里不做过多赘述。

45 步骤 603，针对每个所述模型训练数据，从各个所述特征数据中确定出所述模型训练数据对应的替换数据；基于第二非参数估计方法，确定所述替换数据在各个所述特征数据中的第二分布概率。

针对锅炉 A、锅炉 B 以及锅炉 C 分别对应的每个模型训练数据，从各个特征数据中确

定该模型训练数据对应的替换数据；基于核密度估算方法，确定锅炉 D 对应的 n 个特征数据的概率密度分布函数；针对每个替换数据，基于锅炉 D 对应的 n 个特征数据的概率密度分布函数，确定该替换数据在锅炉 D 中的所有特征数据中的第二分布概率 P_2 。

- 5 步骤 604，针对每个所述模型训练数据，确定所述模型训练数据对应的替换数据的第二分布概率和所述模型训练数据的第一分布概率的分布概率比值；将所述分布概率比值确定为确定所述模型训练数据对应的误差权重。

针对每个模型训练数据，将模型训练数据对应的替换数据的第二分布概率 P_2 ，与模型训练数据的第一分布概率 P_1 的比值 P_2/P_1 ，将 P_2/P_1 确定为该模型训练数据对应的误差权重 W 。

- 10 步骤 605，根据各个所述模型训练数据、各个所述模型训练数据分别对应的设备状态标签以及各个所述模型训练数据分别对应的误差权重进行模型训练，以确定设备状态预测模型，所述设备状态预测模型用于所述目标设备的设备状态预测。

- 具体地，以锅炉 A、锅炉 B 以及锅炉 C 分别对应的 N 个模型训练数据以及每个模型训练数据分别对应的故障标签进行设备状态预测模型的训练，基于每个模型训练数据分别对应的误差权重 W 调整模型参数，最终得到设备状态预测模型。

基于与本发明方法实施例提供一种设备故障预测方法相同的构思，请参考图 7。图 7 为本发明一实施例提供一种设备故障预测装置的结构示意图。如图 7 所示，该设备状态故障预测装置，包括：

- 20 信息确定模块 701，用于确定目标设备的特征信息和目标设备对应的检测点数据信息；
概率模型确定模块 702，用于基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据概率分布模型和具有非共享数据的检测点数据的概率分布模型；

权重确定模块 703，用于根据所述特征数据概率分布模型和所述检测点数据的概率分布模型，确定所述非共享数据的权重；

- 25 模型建立模块 704，用于根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签，建立联邦学习模型；

预测模块 705，用于根据所述联邦学习模型进行所述目标设备的设备故障预测。

- 30 本发明的一个实施例中，所述概率模型确定模块 702，包括：数据确定单元、第一概率分布模型确定单元、第二概率分布模型确定单元；其中，数据确定单元，用于基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据和具有非共享数据的检测点数据；第一概率分布模型确定单元，用于基于所述目标设备的特征信息和目标设备对应的检测点数据信息，确定目标设备的特征数据和具有非共享数据的检测点数据；第二概率分布模型确定单元，用于根据检测点数据参数模型计算所述检测点数据的数据分布，并将确定好参数的检测点数据参数模型确定为所述检测点数据的概率分布模型。

- 35 该实施例中，所述特征数据参数模型和/或检测点数据参数模型包括混合高斯模型。

- 40 本发明的一个实施例中，所述权重确定模块 703，包括：第一分布概率确定单元、第二分布概率单元以及权重确定单元；其中，所述第一分布概率确定单元，用于根据所述特征数据概率分布模型，确定所述非共享数据的目标设备分布概率；所述第二分布概率单元，用于根据所述检测点数据的概率分布模型，确定所述非共享数据的检测点分布概率；所述权重确定单元，用于将所述非共享数据的目标设备分布概率和所述非共享数据的检测点分布概率的比值，确定为所述非共享数据的权重。

- 45 本发明的一个实施例中，所述模型建立模块 704，包括：局部模型确定单元以及模型建立单元；所述局部模型确定单元，用于根据所述非共享数据、所述非共享数据的权重和所述非共享数据对应的设备故障标签，确定检测点局部模型；模型建立单元，用于根据至少两个所述检测点的检测点局部模型，建立联邦学习模型。

本发明的一个实施例中，所述模型建立单元，用于基于联合学习算法，对至少两个所述

检测点的检测点局部模型进行反复迭代，并基于至少两个所述检测点的迭代好的检测点局部模型，建立联邦学习模型。

本发明的一个实施例中，所述检测点局部模型包括神经网络模型或回归模型。

图 8 为本发明一实施例提供的一种设备故障预测装置的结构示意图。如图 8 所示，该设备故障预测装置包括：数据获取模块 801、权重计算模块 802、局部模型训练模块 803、联合模型建立模块 804 和故障预测模块 805，其中，

所述数据获取模块 801，用于根据目标设备属性，获取用于对目标设备建立预测模型的训练数据集，其中所述数据集中的样本数据为共享数据；

所述权重计算模块 802，用于计算所述训练数据集中每条样本数据的权重；

所述局部模型训练模块 803，用于利用所述权重训练得到目标设备的故障预测局部模型；

所述联合模型建立模块 804，用于基于所述故障预测局部模型与联合学习算法，建立联合模型；

所述故障预测模块 805，用于根据所述联合模型对所述目标设备进行故障预测。

在本发明一个实施例中，所述训练数据集中的所述样本数据包含目标设备的特征数据和样本设备的特征数据、样本设备的故障数据；所述样本设备为与所述目标设备相关或相似的设备。

在本发明一个实施例中，所述权重计算模块包括数据标记单元、数据分类单元和数据计算单元，其中，针对每台所述样本设备，所述数据标记单元，用于将所述数据获取模块获取的所述样本设备的特征数据与所述目标设备的特征数据进行区分；所述数据分类单元，用于根据所述数据标记单元区分后的特征数据，训练分类模型；所述数据计算单元，根据所述数据分类单元训练后的分类模型，计算所述样本设备每条特征数据的权重。

在本发明一个实施例中，针对每台所述样本设备：所述数据标记单元，具体用于将所述数据获取模块获取的所述样本设备的特征数据标记为第一数据，将所述数据获取模块获取的所述目标设备的特征数据标记为第二数据；所述数据分类单元，具体用于根据所述数据标记单元标记的所述第一数据和所述第二数据，训练分类模型，所述分类模型为基于联合学习的分类模型；所述数据计算单元，具体用于根据所述数据分类单元训练后的分类模型，计算所述样本设备每条特征数据的权重，所述权重的计算公式为：

$$\omega_i = \frac{P_{2i}}{P_{1i}}$$

其中， ω_i 为所述第一数据中第 i 条数据的权重， P_{1i} 为所述第 i 条数据属于所述样本设备的概率， P_{2i} 为所述第 i 条数据属于所述目标设备的概率。

在本发明一个实施例中，所述局部模型训练模块具体用于根据每台所述样本设备的特征数据、所述样本设备每条特征数据的权重和每台所述样本设备的故障数据，使用神经网络分别在所述样本设备有权重的所述训练数据集上训练得到所述目标设备的故障预测局部模型。

在本发明一个实施例中，所述联合模型建立模块具体用于根据所述故障预测局部模型，使用联合学习算法反复迭代，获得所述样本设备在所述训练数据集上关于所述目标设备的联合模型。

在本发明一个实施例中，所述目标设备和各个所述样本设备均为物联网中的边缘节点，所述目标设备的特征数据不对其他所述样本设备暴露，每个所述样本设备的特征数据和故障数据不对其他所述样本设备和所述目标设备暴露。

基于与本发明方法实施例相同的构思，请参考图 9。图 9 为本发明一实施例提供的一种设备状态预测装置的结构示意图。如图 9 所示，该设备状态预测装置，包括：

获取模块 901，用于获取目标设备对应的至少两个特征数据、至少两个参考设备分别对应的至少两个模型训练数据以及各个所述模型训练数据分别对应的设备状态标签，所述目标设备和所述参考设备的设备种类相同；

权重确定模块 902，用于根据各个所述特征数据，确定所述模型训练数据对应的误差权重；

训练模块 903，用于根据各个所述模型训练数据、各个所述模型训练数据分别对应的设备状态标签以及各个所述模型训练数据分别对应的误差权重进行模型训练，以确定设备状态预测模型，所述设备状态预测模型用于所述目标设备的设备状态预测。

在一个实施例中，所述权重确定模块 902，包括：替换数据确定单元、分布信息确定单元及重要程度确定单元；其中，所述替换数据确定单元，用于从各个所述特征数据中确定出所述模型训练数据对应的替换数据；所述分布信息确定单元，用于确定所述模型训练数据在其所属的参考设备对应的各个所述模型训练数据中的第一分布信息以及所述替换数据在各个所述特征数据中的第二分布信息；所述重要程度确定单元，用于根据所述模型训练数据的第一分布信息以及所述模型训练数据对应的替换数据的第二分布信息，确定所述模型训练数据对应的误差权重。

在一个实施例中，所述分布信息确定单元，包括：第一信息确定子单元以及第二信息确定子单元；其中，所述第一信息确定子单元，用于基于所述模型训练数据在其所属的参考设备对应的第一非参数估计方法，确定所述模型训练数据在其所属的参考设备对应的各个所述模型训练数据中的第一分布概率，将所述第一分布概率确定为第一分布信息；所述第二信息确定子单元，用于基于第二非参数估计方法，确定所述替换数据在各个所述特征数据中的第二分布概率，将所述第二分布概率确定为第二分布信息。

在一个实施例中，所述第一非参数估计方法和所述第二非参数估计方法均为核密度估计算法。

在一个实施例中，所述权重确定模块 302，用于针对每个所述模型训练数据，确定所述模型训练数据对应的替换数据的第二分布概率和所述模型训练数据的第一分布概率的分布概率比值，将所述分布概率比值确定为确定所述模型训练数据对应的误差权重。

在一个实施例中，所述方法应用到所述目标设备对应的边缘服务器或本地服务器。

在一个实施例中，所述训练模块 903，包括：训练单元、判断单元以及调整单元；其中，所述训练单元，用于根据将所述模型训练数据代入待预测模型中的预测结果以及所述模型训练数据对应的设备状态标签，确定所述模型训练数据对应的第一误差，根据所述模型训练数据对应的第一误差以及误差权重，以确定所述模型训练数据对应的第二误差；所述判断单元，用于判断是否满足迭代次数或者各个所述模型训练数据分别对应的第二误差是否满足预设条件，如果是，则将所述待预测模型确定为设备状态预测模型，如果否，则触发所述调整单元；所述调整单元，用于根据各个所述模型训练数据分别对应的第二误差，对所述待预测模型中的模型参数进行调整，以确定调整后的模型参数，并将所述待预测模型中的模型参数替换为所述调整后的模型参数，触发所述训练单元。

上述装置内的各模块和单元之间的信息交互、执行过程等内容，由于与本发明方法实施例基于同一构思，具体内容可参见本发明方法实施例中的叙述，此处不再赘述。

图 10 是本发明实施例提供的一种电子设备的结构示意图。在硬件层面，该电子设备包括处理器 1001 以及存储有执行指令的存储器 1002，可选地还包括内部总线 1003 及网络接口 1004。其中，存储器 1002 可能包含内存 10021，例如高速随机存取存储器（Random-Access Memory，RAM），也可能还包括非易失性存储器 10022（Non-Volatile Memory），例如至少 1 个磁盘存储器等；处理器 1001、网络接口 1004 和存储器 1002 可以通过内部总线 1003 相互连接，该内部总线 1003 可以是 ISA（Industry Standard Architecture，工业标准体系结构）总线、PCI（Peripheral Component Interconnect，外设部件互连标准）总线或 EISA（Extended Industry Standard Architecture，扩展工业标准结构）总线等；内部总线 1003 可以分为地址总线、数据总线、控制总线等，为便于表示，图 10 中仅用一个双向箭头表示，但并不表示仅有一根总线或一种类型的总线。当然，该电子设备还可能包括其他业务所需要的硬件。当处理器 1001 执行存储器 1002 存储的执行指令时，处理器 1001 执行本发明任意一个实施例中

的方法。

在一种可能实现的方式中，处理器从非易失性存储器中读取对应的执行指令到内存中然后运行，也可从其它设备上获取相应的执行指令，以在逻辑层面上形成一种设备故障预测装置。处理器执行存储器所存放的执行指令，以通过执行的执行指令实现本发明任一实施例中提供的一种设备故障预测方法。

5 处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器（Central Processing Unit，CPU）、网络处理器（Network Processor，NP）等；还可以是数字信号处理器（Digital Signal Processor，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现场可编程门阵列（Field-Programmable Gate Array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

10 本发明实施例还提供了一种计算机可读存储介质，包括执行指令，当电子设备的处理器执行执行指令时，所述处理器执行本发明任意一个实施例中提供的方法。该电子设备具体可以是如图 10 所示的电子设备；执行指令是一种设备故障预测方法所对应计算机程序。

本领域内的技术人员应明白，本发明的实施例可提供为方法或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例，或软件和硬件相结合的形式。

20 本发明中的多个实施例均采用递进的方式描述，多个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

25 还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

30 以上所述仅为本发明的实施例而已，并不用于限制本发明。对于本领域技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本发明的权利要求范围之内。