

(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
2022 年 6 月 30 日 (30.06.2022)



(10) 国际公布号
WO 2022/134809 A1

(51) 国际专利分类号:
G06N 20/00 (2019.01)

(21) 国际申请号: PCT/CN2021/125440

(22) 国际申请日: 2021 年 10 月 21 日 (21.10.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011553263.7 2020 年 12 月 24 日 (24.12.2020) CN

(71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(72) 发明人: 张山 (ZHANG, Shan); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(74) 代理人: 深圳市赛恩倍吉知识产权代理有限公司(SHENZHEN SCIENBIZIP INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国广东省深圳市龙华新区龙观东路 83 号荣群大厦 9 楼, Guangdong 518109 (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, IT, 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,

(54) Title: MODEL TRAINING PROCESSING METHOD AND APPARATUS, COMPUTER DEVICE, AND MEDIUM

(54) 发明名称: 模型训练处理方法、装置、计算机设备及介质

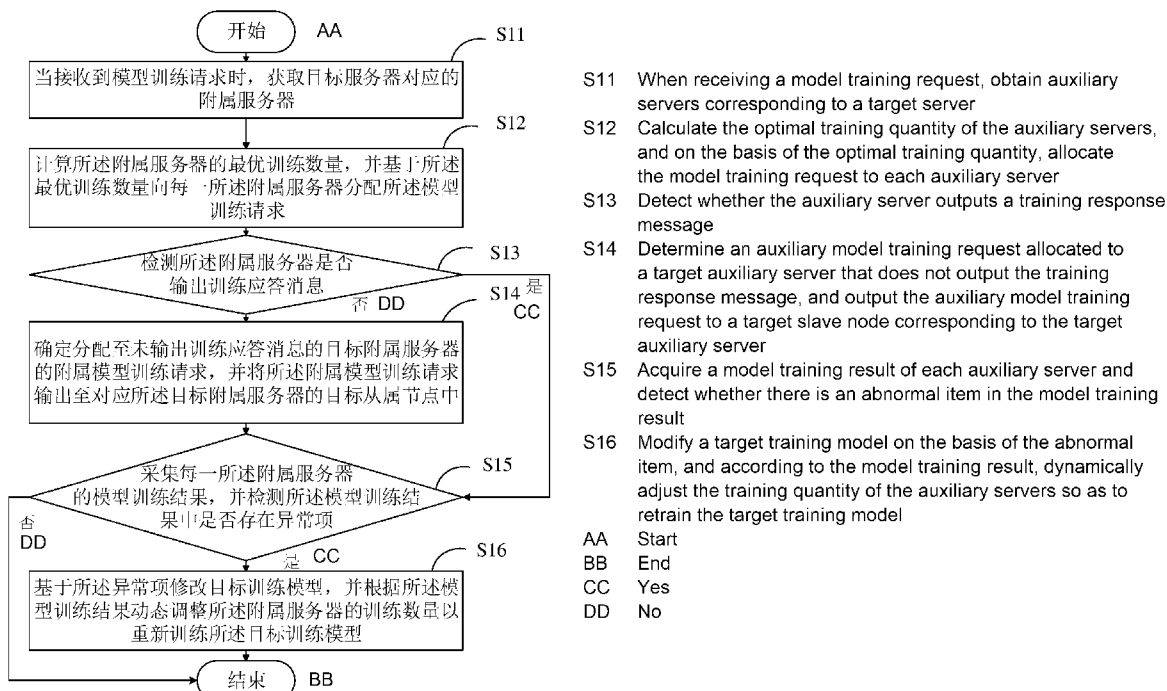


图 1

(57) Abstract: The present application relates to data processing technology, and provides a model training processing method and apparatus, a computer device, and a storage medium. The method comprises: obtaining auxiliary servers corresponding to a target server; calculating the optimal training quantity of the auxiliary servers, and allocating a model training request to each auxiliary server; detecting whether the auxiliary server outputs a training response message; when the detection result is no, determining an auxiliary model training request, and outputting the auxiliary model training request to a corresponding target slave node; acquiring a model training result of each auxiliary server and detecting whether there is an abnormal item in the model training result; when the detection result is no, determining a target training model on the basis of the abnormal item, and according to the model training result, dynamically adjusting the training quantity of the auxiliary servers so as to retrain the target training model.

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,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, 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))。

training result of each auxiliary server, and detecting whether there is an abnormal item in the model training result; and when the detection result is yes, modifying a target training model on the basis of the abnormal item, and according to the model training result, dynamically adjusting the training quantity of the auxiliary servers so as to retrain the target training model. The present application can increase the model training efficiency, and promotes the construction of smart healthcare and smart cities.

(57) 摘要: 本申请涉及数据处理技术, 提供一种模型训练处理方法、装置、计算机设备与存储介质, 包括: 获取目标服务器对应的附属服务器; 计算附属服务器的最优训练数量, 并向每一附属服务器分配模型训练请求; 检测附属服务器是否输出训练应答消息; 当检测结果为否时, 确定附属模型训练请求, 并将附属模型训练请求输出至对应目标从属节点中; 采集每一附属服务器的模型训练结果, 并检测模型训练结果中是否存在异常项; 当检测结果为是时, 基于异常项修改目标训练模型, 并根据模型训练结果动态调整附属服务器的训练数量以重新训练目标训练模型。本申请能够提高模型训练效率, 促进智慧医疗及智慧城市的建设。

模型训练处理方法、装置、计算机设备及介质

本申请要求于 2020 年 12 月 24 日提交中国专利局，申请号为 202011553263.7 发明名称为“模型训练处理方法、装置、计算机设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及数据处理技术领域，尤其涉及一种模型训练处理方法、装置、计算机设备及介质。

背景技术

深度神经网络已成功应用于很多领域，包括图像识别、纹理分类、语音识别等领域。近年来，深度神经网络由于采用更深的网络架构和更大的训练样本集来进行训练，所以在性能上有了显著的提升。

针对深度神经网络训练的情况，在实现本申请的过程中，发明人发现现有技术至少存在如下问题：深度神经网络学习需要花费太高的时间成本，深度神经网络学习动辄就几个小时甚至花费几天时间来进行，一旦算法出现问题，就要重新训练深度神经网络学习模型，模型训练的效率极低。

因此，有必要提供一种模型训练处理方法，能够减少模型训练所需时间，提高模型训练效率。

发明内容

鉴于以上内容，有必要提出一种模型训练处理方法、模型训练处理装置、计算机设备及介质，能够减少模型训练所需时间，提高模型训练效率。

本申请实施例第一方面提供一种模型训练处理方法，所述模型训练处理方法包括：

当接收到模型训练请求时，获取目标服务器对应的附属服务器；

计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求；

检测所述附属服务器是否输出训练应答消息；

当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

本申请实施例第二方面提供的上述一种模型训练处理装置，所述模型训练处理装置包括：请求接收模块，用于当接收到模型训练请求时，获取目标服务器对应的附属服务器；

数量计算模块，用于计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求；

消息检测模块，用于检测所述附属服务器是否输出训练应答消息；

节点确定模块，用于当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

异常检测模块，用于采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

数量调整模块，用于当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

本申请实施例第三方面还提供一种计算机设备，所述计算机设备包括处理器，所述处理器用于执行存储器中存储的计算机可读指令以实现以下步骤：

当接收到模型训练请求时，获取目标服务器对应的附属服务器；

计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求；

检测所述附属服务器是否输出训练应答消息；

当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

本申请实施例第四方面还提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现以下步骤：

当接收到模型训练请求时，获取目标服务器对应的附属服务器；

计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求；

检测所述附属服务器是否输出训练应答消息；

当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

本申请能够减少模型训练所需时间，提高模型训练效率。

附图说明

图 1 是本申请实施例一提供的模型训练处理方法的流程图。

图 2 是本申请实施例二提供的模型训练处理装置的结构图。

图 3 是本申请实施例三提供的计算机设备的结构示意图。

如下具体实施方式将结合上述附图进一步说明本申请。

具体实施方式

为了能够更清楚地理解本申请的上述目的、特征和优点，下面结合附图和具体实施例对本申请进行详细描述。需要说明的是，在不冲突的情况下，本申请的实施例及实施例中的特征可以相互组合。

在下面的描述中阐述了很多具体细节以便于充分理解本申请，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

除非另有定义，本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同。本文中在本申请的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本申请。

本申请实施例提供的模型训练处理方法由计算机设备执行，相应地，模型训练处理装置运行于计算机设备中。

图 1 是本申请实施例一提供的模型训练处理方法的流程图。如图 1 所示，所述模型训练处理方法可以包括如下步骤，根据不同的需求，该流程图中步骤的顺序可以改变，某些可以省略：

S11、当接收到模型训练请求时，获取目标服务器对应的附属服务器。

在本申请的至少一实施例中，所述模型训练请求是指用于请求训练目标训练模型的请求，所述模型训练请求是向目标应用发出的请求，用于开启所述目标训练模型的训练，

所述目标应用是指包含所述目标训练模型应用，所述应用在接收到模型训练请求时，将所述模型训练请求转发至目标服务器集群中，通过所述目标服务器集群中的多个服务器并行执行所述模型训练请求。所述目标服务器集群包括目标服务器以及对应所述目标服务器的附属服务器。所述目标服务器是指接收所述目标应用转发模型训练请求的服务器，所述附属服务器是指与所述目标服务器关联的服务器。可选地，所述当接收到模型训练请求时，获取目标服务器对应的附属服务器包括：解析所述模型训练请求，得到对应所述模型训练请求的目标路由信息；根据所述目标路由信息确定目标应用；获取对应所述目标应用的目标服务器集群；获取所述目标服务器集群中的目标服务器与附属服务器。

其中，每一所述模型训练请求均包含目标路由信息，所述目标路由信息用于指示对应目标训练模型的目标应用的路由信息，根据所述目标路由信息能够得到对应所述模型训练请求的目标应用。所述目标应用与目标服务器集群相关联，根据目标应用能够得到对应的目标服务器集群。所述目标服务器集群是指包含一个目标服务器以及多个附属服务器的集群，所述附属服务器的数量可以为1个，也可以为多个，在此不作限制。

在本申请的至少一实施例中，所述附属服务器的数量可以根据所述模型训练请求中携带的训练数量确定，所述训练数量越大，其对应的所述附属服务器的数量也越大。可选地，所述获取目标服务器对应的附属服务器还包括：解析所述模型训练请求，得到对应目标训练模型的训练数量；确定所述训练数量对应的训练数量级；根据所述训练数量级确定附属服务器的数量。

其中，所述模型训练请求不仅包含目标路由信息，还包括对应目标训练模型的训练数量等信息，在此不作限制。所述训练数量是指训练数据的数量，所述训练数量可以根据目标训练模型的实际情况进行设定。在一实施例中，所述训练数量与所述训练数量级间存在预先设置的第一映射关系，所述训练数量级与所述附属服务器的数量间存在预先设置的第二映射关系，通过所述训练数量遍历所述第一映射关系，能够确定对应所述训练数量的训练数量级，所述训练数量级可以为低级、中级与高级。通过所述训练数量级遍历所述第二映射关系，能够确定对应所述训练数量级的附属服务器的数量，避免调用过多或过少的附属服务器造成的资源浪费或模型训练时间长的的问题。

S12、计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求。

在本申请的至少一实施例中，不同附属服务器的数据处理能力可能相同，也可能不相同。当每个附属服务器的数据处理能力均相同时，可以将所述模拟训练请求携带的训练数量进行等量拆分，得到附属模型训练请求，并将所述附属模型训练请求分配至所述附属服务器。当不同附属服务器的数据处理能力不相同，由于各个服务器的数据处理能力相同，导致相同或者相近的时间段内不同附属服务器处理训练数据的数量不同，本申请根据附属服务器的不同数据处理能力，为每一所述附属服务器分配最优训练数量，

使得数据处理能力高的附属服务器处理较多的训练数据，而数据处理能力少的附属服务器处理较少的训练数据，从而附属服务器均能够在相同或相近的时间段处理完各自的训练数量，减少模型训练整体所需的时间，提高模型训练的效率。

其中，所述附属服务器的数据处理能力可根据服务器的负载等级确定，所述负载等级越高，其对应的数据处理能力越高，所述负载等级与所述数据处理能力间存在映射关系，通过查询该映射关系，能够得到所述负载等级对应的数据处理能力。所述负载等级可通过服务器的 CPU 数据、内存数据与网络负载数据这 3 个指标确定，所述 CPU 数据可以是指 CPU 负载参数，所述内存数据可以是指内存占用参数，所述网络负载数据可以包括流量参数，在此不做限制。在一实施例中，通过获取上述 3 个指标，并将其输入至预先训练的负载等级确定模型中进行处理，得到对应服务器的负载等级。

可选地，所述计算所述附属服务器的最优训练数量包括：获取对应每一附属服务器的数据处理能力；获取所述模型训练请求携带的对应目标训练模型的训练数量；根据所述数据处理能力与所述训练数量计算每一所述附属服务器的最优训练数量。其中，所述模型训练请求中携带的训练数量可以是指训练数据的总量，所述数据处理能力可以是指单位时间内服务器处理模型训练数据的数量，通过配置相同或相近的训练时间，能够得到分配至各个服务器的最优训练数量。

在本申请的至少一实施例中，当每一所述附属服务器的最优训练数量确定后，开始对所述模型训练请求进行分配处理，将所述模型训练请求拆分为多个附属模型训练请求，每一所述附属模型训练请求与所述附属服务器相对应，用于请求调用所述附属服务器对训练数据进行训练处理。可选地，所述基于所述最优训练数量分配所述模型训练请求包括：获取所述附属服务器的数量与每一所述附属服务器的最优训练数量；基于所述附属服务器的数量与所述最优训练数量拆分所述模型训练请求，得到若干个附属模型训练请求；将所述附属模型训练请求分配至对应的附属服务器。

其中，每一所述附属模型训练请求中均携带有该附属服务器的标识以及该附属服务器应训练的训练数量，通过解析所述附属模型训练请求即可将所述附属模型训练请求分配至对应的附属服务器中。

S13、检测所述附属服务器是否输出训练应答消息，当检测结果为所述附属服务器未输出训练应答消息时，执行步骤 S14。

在本申请的至少一实施例中，每一所述附属服务器在开始训练前，均会向目标服务器输出训练应答消息，通过检查所述附属服务器是否输出训练应答消息，能够确定所述附属服务器是否正常训练模型数据。当检测结果为所述附属服务器输出训练应答消息时，确定所述附属服务器训练进程正常；当检测结果为所述附属服务器未输出训练应答消息时，确定所述附属服务器训练进程异常。

可选地，所述检测所述附属服务器是否输出训练应答消息包括：获取对应各个所述

附属服务器的运行日志；检测所述运行日志中是否存在每一所述附属服务器对应的目标应答数据；当检测结果为所述运行日志中存在每一所述附属服务器对应的目标应答数据时，确定所述附属服务器输出训练应答消息；当检测结果为所述运行日志中不存在所述附属服务器对应的目标应答数据时，确定所述附属服务器未输出训练应答消息。

其中，各个所述附属服务器的运行日志可以存储于区块链的目标节点中，各个附属服务器将运行过程中产生的训练应答消息、运行异常消息以及运行结束消息等数据存储至所述目标节点中，通过查询所述运行日志，能够得到各个附属服务器的运行进程。此外，所述运行日志内还可以记录有各个服务器的地址信息、端口信息、从属节点信息、从属节点的端口信息（也即从端口信息）、节点类型（所述节点类型可以包括主任务与附属任务）以及附属模型训练请求信息等，在此不做限制。

在本申请的至少一实施例中，当检测结果为所述附属服务器均输出训练应答消息时，则执行步骤 S15，用于采集每一所述附属服务器的模型训练结果。

S14、确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中。

在本申请的至少一实施例中，每一所述附属服务器均设有从属节点，所述从属节点可配置与所述附属服务器相对应的模型训练处理逻辑，当所述目标附属服务器未输出训练应答消息时，说明此时所述目标附属服务器运行进程存在异常。本申请通过调用运行进程存在异常的目标附属服务器对应的目标从属节点执行相应的模型训练请求，保证模型训练请求能够正常执行，节省模型训练所需的时间，提高模型训练的效率。

可选地，所述确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中包括：获取所述目标附属服务器的从属节点集；从所述从属节点集中确定目标从属节点；将所述附属模型训练请求输出至所述目标从属节点，由所述目标从属节点执行模型训练。

其中，所述目标附属服务器对应的从属节点的数量可以为 1 个，也可以为多个。当所述目标附属服务器对应的从属节点的数量为多个时，所述方法还包括：获取所述目标附属服务器的从属节点集；确定所述从属节点集中每一从属节点对应的执行逻辑；检测是否存在与目标附属服务器的执行逻辑一致的目标执行逻辑；当检测结果为存在与所述目标附属服务器的执行逻辑一致的目标执行逻辑时，确定对应所述目标执行逻辑的从属节点为目标从属节点。

可选地，所述从属节点大多用于提供查询问答匹配逻辑，可配置一个从属节点用于配置任务处理逻辑，用于在附属服务器无法正常执行模型训练任务时，调用该附属服务器对应的从属节点执行模型训练任务。在一实施例中，针对每一附属服务器，均预先配置目标从属节点，用于执行任务处理逻辑。在其他实施例中，无需预先配置从属节点的任务处理逻辑，当检测到附属服务器的任务执行进程异常时，获取对应所述附属服务器

的目标执行逻辑信息，将所述目标执行逻辑信息输出至所述从属节点集中的任一从属节点中，并根据所述目标执行逻辑信息配置该从属节点，使得该从属节点具有任务处理逻辑。本申请在检测到附属服务器的任务执行异常时，根据目标执行逻辑信息实时配置从属节点，避免预先配置导致的资源浪费的情况，提高开发效率。

其中，所述从属节点预先设有相关配置信息，用于执行对应查询问答匹配逻辑，所述根据所述目标执行逻辑信息配置该从属节点也即将所述目标执行逻辑信息存储至对应从属节点的配置信息中，并调用所述目标执行逻辑信息使得所述从属节点具备任务处理能力。

S15、采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项，当检测结果为所述模型训练结果中存在异常项时，执行步骤 S16。

在本申请的至少一实施例中，每一所述附属服务器在模型训练过程中，均会向所述区块链的目标节点的运行日志中输出相关训练数据，当每一所述附属服务器的模型训练结束后，会向所述目标节点的运行日志中输出模型训练结果。本申请除了通过查询所述运行日志得到每一所述附属服务器的模型训练结果外，所述方法还包括：获取所述目标服务器与所述附属服务器的关联关系；根据所述关联关系确定可视化的模型训练逻辑图，其中，每一服务器对应所述模型训练逻辑图中的目标按钮；在所述可视化的模型训练逻辑图中，获取每一所述附属服务器的模型训练进度与训练速率；根据所述训练进度与所述训练速率确定每一所述目标按钮的实化显示面积与实化显示速率。

其中，所述模型训练逻辑图对应各个服务器的目标按钮，每一按钮均对应一附属服务器，所述目标按钮与所述附属服务器间可通过预设标识进行关联，通过查看所述预设标识，即可确定每一所述目标按钮关联的服务器，所述预设标识可以为数字标识、字母标识与颜色标识，在此不做限制。本申请通过对每一所述目标按钮进行虚化与实化显示，能够直观地显示每一所述附属服务器的训练进度与训练速率。当所述目标按钮的实化显示面积为整个目标按钮面积时，确定所述服务器训练完成；所述目标按钮的实化显示速率越快，其对应服务器的训练速率越快。

在本申请的至少一实施例中，在可视化的模型逻辑训练图中，除了可以通过虚化显示的方式直观地表示附属服务器的模型训练进度与模型训练速率，所述方法还包括：检测所述附属服务器的运行过程中是否出现训练异常项；当检测结果为所述附属服务器的运行过程中出现训练异常项时，获取所述异常项的数量；根据所述异常项的数量在可视化的模型训练逻辑图的目标按钮中设置预设标识。

其中，所述异常项的数量不同，所述预设标识也不相同，例如，所述预设标识为颜色标识时，当所述异常项的数量越多时，所述预设标识的越接近红色，在此不作限制。通过在可视化的模型训练逻辑图中显示异常项的预设标识，能够直观地看出附属服务器在模型训练过程中是否出现异常项以及出现的异常项的数量，并及时对异常情况进行查

看。

S16、基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

在本申请的至少一实施例中，当检测结果为所述模型训练结果中存在异常项时，开发人员可基于所述异常项修改目标训练模型，以使得修改好的目标训练模型能够解决上述异常项。在实际训练过程中，当开发人员根据所述异常项修改好目标训练模型后，还需将所述目标训练模型按照训练数据重新进行训练，以确保目标训练模型在正式使用时无其他异常项。所述模型训练结果中还包含各个附属服务器的训练时长，本申请根据所述训练时长动态微调调整所述附属服务器的训练数据数量，进一步节省模型训练的时间，提高模型训练效率。

可选地，所述根据所述模型训练结果动态调整所述附属服务器的训练数量包括：获取并解析每一所述附属服务器的模型训练结果，得到对应每一所述附属服务器的实际模型训练时长；检测所述实际模型训练时长是否满足目标调整条件；当检测结果为所述实际模型训练时长满足目标调整条件时，确定满足目标调整条件的所述目标附属服务器；动态调整所述目标附属服务器的训练数量。

其中，所述目标调整条件是指所述实际模型训练时长中存在异常时长，举例来说，存在 5 个附属服务器，分别为附属服务器 1-5，其中，附属服务器 1-4 彼此间的实际模型训练时长相差时间间隔均在预设时间间隔内，而附属服务器 5 的实际模型训练时长与其余 4 个附属服务器中一个或多个的实际模型训练时长的时间间隔超过预设时间间隔，则附属服务器 5 的实际模型训练时长满足目标调整条件，需要对所述附属服务器 5 的训练数量进行动态调整，以使得所述附属服务器 5 的训练时长与其他几个附属服务器的训练时长的时间间隔在预设时间间隔内，进一步从整体上减少模型训练所需时间，提高模型训练的效率。

可选地，所述动态调整所述附属服务器的训练数量包括：确定实际模型训练时长最长的第一附属服务器与实际模型训练时长最短的第二附属服务器；计算所述第一附属服务器与所述第二附属服务器的训练时长间隔；获取所述第一附属服务器的第一训练数量以及所述第二附属服务器的第二训练数量；根据所述训练时长间隔调整所述第一训练数量与所述第二训练数量。

本申请提供的上述模型训练处理方法，将单台服务器单线程的训练任务处理扩展为多台服务器并行处理，使得模型训练时间大幅降低，提高模型训练效率；且本申请针对每台附属服务器的最优训练数量分配模型训练请求，使得每台附属服务器的训练时间间隔差距较小，使得模型训练时间降低，训练时间能够缩短至 30 分钟以内，提高模型训练效率；此外，本申请基于异常项修改目标训练模型后，根据此前的模型训练结果动态调整附属服务器的训练数量，进一步优化模型训练过程，提高模型训练效率；此外，本申

请通过对集群从属节点的配置，使得从属节点支持逻辑匹配查询，增加了系统的并发查询能力，加强了系统的可靠性。本申请可应用于智慧政务、智慧交通等智慧城市的各个功能模块中，比如智慧政务的模型训练处理模块等，能够促进智慧城市的快速发展。

图 2 是本申请实施例二提供的模型训练处理装置的结构图。

在一些实施例中，所述模型训练处理装置 20 可以包括多个由计算机程序段所组成的功能模块。所述模型训练处理装置 20 中的各个程序段的计算机程序可以存储于计算机设备的存储器中，并由至少一个处理器所执行，以执行（详见图 1 描述）模型训练处理的功能。

本实施例中，所述模型训练处理装置 20 根据其所执行的功能，可以被划分为多个功能模块。所述功能模块可以包括：请求接收模块 201、数量计算模块 202、消息检测模块 203、节点确定模块 204、异常检测模块 205 以及数量调整模块 206。本申请所称的模块是指一种能够被至少一个处理器所执行并且能够完成固定功能的一系列计算机程序段，其存储在存储器中。在本实施例中，关于各模块的功能将在后续的实施例中详述。

请求接收模块 201 可以用于当接收到模型训练请求时，获取目标服务器对应的附属服务器。

数量计算模块 202 可以用于计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求。

消息检测模块 203 可以用于检测所述附属服务器是否输出训练应答消息。

节点确定模块 204 可以用于当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中。

异常检测模块 205 可以用于采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项。

数量调整模块 206 可以用于当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

参阅图 3 所示，为本申请实施例三提供的计算机设备的结构示意图。在本申请较佳实施例中，所述计算机设备 3 包括存储器 31、至少一个处理器 32、至少一条通信总线 33 及收发器 34。

本领域技术人员应该了解，图 3 示出的计算机设备的结构并不构成本申请实施例的限定，既可以是总线型结构，也可以是星形结构，所述计算机设备 3 还可以包括比图示更多或更少的其他硬件或者软件，或者不同的部件布置。

在一些实施例中，所述计算机设备 3 是一种能够按照事先设定或存储的指令，自动进行数值计算和/或信息处理的设备，其硬件包括但不限于微处理器、专用集成电路、可编程门阵列、数字处理器及嵌入式设备等。所述计算机设备 3 还可包括客户设备，所述客户设备包括

但不限于任何一种可与客户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互的电子产品，例如，个人计算机、平板电脑、智能手机、数码相机等。

需要说明的是，所述计算机设备 3 仅为举例，其他现有的或今后可能出现的电子产品如可适应于本申请，也应包含在本申请的保护范围以内，并以引用方式包含于此。

在一些实施例中，所述存储器 31 中存储有计算机程序，所述计算机程序被所述至少一个处理器 32 执行时实现如所述的模型训练处理方法中的全部或者部分步骤。示例性的，所述计算机程序可以被分割成一个或多个模块/单元，所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机可读指令段，该指令段用于描述所述计算机程序在所述计算机设备中的执行过程。例如，图 2 中所述的各个模块是存储在所述存储器 31 中的计算机程序，并由所述至少一个处理器 32 所执行，从而实现所述各个模块的功能以达到基于大数据的信息处理的目的。所述存储器 31 包括只读存储器（Read-Only Memory, ROM）、可编程只读存储器（Programmable Read-Only Memory, PROM）、可擦除可编程只读存储器（Erasable Programmable Read-Only Memory, EPROM）、一次可编程只读存储器（One-time Programmable Read-Only Memory, OTPROM）、电子擦除式可复写只读存储器（Electrically-Erasable Programmable Read-Only Memory, EEPROM）、只读光盘（Compact Disc Read-Only Memory, CD-ROM）或其他光盘存储器、磁盘存储器、磁带存储器、或者能够用于携带或存储数据的计算机可读的任何其他介质。

进一步地，所述计算机可读存储介质可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序等；存储数据区可存储根据区块链节点的使用所创建的数据等。所述计算机可读存储介质可以是非易失性，也可以是易失性。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链（Blockchain），本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性（防伪）和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

在一些实施例中，所述至少一个处理器 32 是所述计算机设备 3 的控制核心（Control Unit），利用各种接口和线路连接整个计算机设备 3 的各个部件，通过运行或执行存储在所述存储器 31 内的程序或者模块，以及调用存储在所述存储器 31 内的数据，以执行计算机设备 3 的各种功能和处理数据。例如，所述至少一个处理器 32 执行所述存储器中存储的计算机程序时实现本申请实施例中所述的模型训练处理方法的全部或者部分步骤；或者实现模型训练处理装置的全部或者部分功能。所述至少一个处理器 32 可以由集成电路组成，例如可以由单个封装的集成电路所组成，也可以是由多个相同功能或不同功能封装的集成电路所组成，包括一个或者多个中央处理器（Central Processing unit, CPU）、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。

在一些实施例中，所述至少一条通信总线 33 被设置为实现所述存储器 31 以及所述至少一个处理器 32 等之间的连接通信。

尽管未示出，所述计算机设备 3 还可以包括给各个部件供电的电源（比如电池），优选的，电源可以通过电源管理装置与所述至少一个处理器 32 逻辑相连，从而通过电源管理装置实现管理充电、放电、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述计算机设备 3 还可以包括多种传感器、蓝牙模块、Wi-Fi 模块等，在此不再赘述。

上述以软件功能模块的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能模块存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，计算机设备，或者网络设备）或处理器（processor）执行本申请各个实施例所述方法的部分。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，既可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能模块的形式实现。

对于本领域技术人员而言，显然本申请不限于上述示范性实施例的细节，而且在不背离本申请的精神或基本特征的情况下，能够以其他的具体形式实现本申请。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本申请的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外，显然“包括”一词不排除其他单元或，单数不排除复数。说明书中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一，第二等词语用来表示名称，而并不表示任何特定的顺序。

最后应说明的是，以上实施例仅用以说明本申请的技术方案而非限制，尽管参照较佳实施例对本申请进行了详细说明，本领域的普通技术人员应当理解，可以对本申请的技术方案进行修改或等同替换，而不脱离本申请技术方案的精神和范围。

权 利 要 求 书

1. 一种模型训练处理方法, 其中, 所述模型训练处理方法包括:
当接收到模型训练请求时, 获取目标服务器对应的附属服务器;
计算所述附属服务器的最优训练数量, 并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求;
检测所述附属服务器是否输出训练应答消息;
当检测结果为所述附属服务器未输出训练应答消息时, 确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求, 并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中;
采集每一所述附属服务器的模型训练结果, 并检测所述模型训练结果中是否存在异常项;
当检测结果为所述模型训练结果中存在异常项时, 基于所述异常项修改目标训练模型, 并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。
2. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述当接收到模型训练请求时, 获取目标服务器对应的附属服务器包括:
解析所述模型训练请求, 得到对应所述模型训练请求的目标路由信息;
根据所述目标路由信息确定目标应用;
获取对应所述目标应用的目标服务器集群;
获取所述目标服务器集群中的目标服务器与附属服务器。
3. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述获取目标服务器对应的附属服务器还包括:
解析所述模型训练请求, 得到对应目标训练模型的训练数量;
确定所述训练数量对应的训练数量级;
根据所述训练数量级确定附属服务器的数量。
4. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述计算所述附属服务器的最优训练数量包括:
获取对应每一附属服务器的数据处理能力;
获取所述模型训练请求携带的对应目标训练模型的训练数量;
根据所述数据处理能力与所述训练数量计算每一所述附属服务器的最优训练数量。
5. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述基于所述最优训练数量分配所述模型训练请求包括:
获取所述附属服务器的数量与每一所述附属服务器的最优训练数量;
基于所述附属服务器的数量与所述最优训练数量拆分所述模型训练请求, 得到若干个附属模型训练请求;

将所述附属模型训练请求分配至对应的附属服务器。

6. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述检测所述附属服务器是否输出训练应答消息包括:

获取对应各个所述附属服务器的运行日志;

检测所述运行日志中是否存在每一所述附属服务器对应的目标应答数据;

当检测结果为所述运行日志中存在每一所述附属服务器对应的目标应答数据时, 确定所述附属服务器输出训练应答消息;

当检测结果为所述运行日志中不存在所述附属服务器对应的目标应答数据时, 确定所述附属服务器未输出训练应答消息。

7. 根据权利要求 1 所述的模型训练处理方法, 其中, 所述方法还包括:

获取所述目标服务器与所述附属服务器的关联关系;

根据所述关联关系确定可视化的模型训练逻辑图, 其中, 每一服务器对应所述模型训练逻辑图中的目标按钮;

在所述可视化的模型训练逻辑图中, 获取每一所述附属服务器的模型训练进度与训练速率;

根据所述训练进度与所述训练速率确定每一所述目标按钮的实化显示面积与实化显示速率。

8. 一种模型训练处理装置, 其中, 所述模型训练处理装置包括:

请求接收模块, 用于当接收到模型训练请求时, 获取目标服务器对应的附属服务器;

数量计算模块, 用于计算所述附属服务器的最优训练数量, 并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求;

消息检测模块, 用于检测所述附属服务器是否输出训练应答消息;

节点确定模块, 用于当检测结果为所述附属服务器未输出训练应答消息时, 确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求, 并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中;

异常检测模块, 用于采集每一所述附属服务器的模型训练结果, 并检测所述模型训练结果中是否存在异常项;

数量调整模块, 用于当检测结果为所述模型训练结果中存在异常项时, 基于所述异常项修改目标训练模型, 并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

9. 一种计算机设备, 其中, 所述计算机设备包括处理器, 所述处理器用于执行存储器中存储的计算机可读指令以实现以下步骤:

当接收到模型训练请求时, 获取目标服务器对应的附属服务器;

计算所述附属服务器的最优训练数量, 并基于所述最优训练数量向每一所述附属服务器

分配所述模型训练请求；

检测所述附属服务器是否输出训练应答消息；

当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

10. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现所述当接收到模型训练请求时，获取目标服务器对应的附属服务器时，包括：

解析所述模型训练请求，得到对应所述模型训练请求的目标路由信息；

根据所述目标路由信息确定目标应用；

获取对应所述目标应用的目标服务器集群；

获取所述目标服务器集群中的目标服务器与附属服务器。

11. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现所述获取目标服务器对应的附属服务器时，还包括：

解析所述模型训练请求，得到对应目标训练模型的训练数量；

确定所述训练数量对应的训练数量级；

根据所述训练数量级确定附属服务器的数量。

12. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现所述计算所述附属服务器的最优训练数量时，包括：

获取对应每一附属服务器的数据处理能力；

获取所述模型训练请求携带的对应目标训练模型的训练数量；

根据所述数据处理能力与所述训练数量计算每一所述附属服务器的最优训练数量。

13. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现所述基于所述最优训练数量分配所述模型训练请求时，包括：

获取所述附属服务器的数量与每一所述附属服务器的最优训练数量；

基于所述附属服务器的数量与所述最优训练数量拆分所述模型训练请求，得到若干个附属模型训练请求；

将所述附属模型训练请求分配至对应的附属服务器。

14. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现所述检测所述附属服务器是否输出训练应答消息时，包括：

获取对应各个所述附属服务器的运行日志；

检测所述运行日志中是否存在每一所述附属服务器对应的目标应答数据；

当检测结果为所述运行日志中存在每一所述附属服务器对应的目标应答数据时，确定所述附属服务器输出训练应答消息；

当检测结果为所述运行日志中不存在所述附属服务器对应的目标应答数据时，确定所述附属服务器未输出训练应答消息。

15. 根据权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令还用以实现以下步骤：

获取所述目标服务器与所述附属服务器的关联关系；

根据所述关联关系确定可视化的模型训练逻辑图，其中，每一服务器对应所述模型训练逻辑图中的目标按钮；

在所述可视化的模型训练逻辑图中，获取每一所述附属服务器的模型训练进度与训练速率；

根据所述训练进度与所述训练速率确定每一所述目标按钮的实化显示面积与实化显示速率。

16. 一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，其中，所述计算机可读指令被处理器执行时实现以下步骤：

当接收到模型训练请求时，获取目标服务器对应的附属服务器；

计算所述附属服务器的最优训练数量，并基于所述最优训练数量向每一所述附属服务器分配所述模型训练请求；

检测所述附属服务器是否输出训练应答消息；

当检测结果为所述附属服务器未输出训练应答消息时，确定分配至未输出训练应答消息的目标附属服务器的附属模型训练请求，并将所述附属模型训练请求输出至对应所述目标附属服务器的目标从属节点中；

采集每一所述附属服务器的模型训练结果，并检测所述模型训练结果中是否存在异常项；

当检测结果为所述模型训练结果中存在异常项时，基于所述异常项修改目标训练模型，并根据所述模型训练结果动态调整所述附属服务器的训练数量以重新训练所述目标训练模型。

17. 根据权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被处理器执行以实现所述当接收到模型训练请求时，获取目标服务器对应的附属服务器时，包括：

解析所述模型训练请求，得到对应所述模型训练请求的目标路由信息；

根据所述目标路由信息确定目标应用；

获取对应所述目标应用的目标服务器集群；

获取所述目标服务器集群中的目标服务器与附属服务器。

18. 根据权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被处理器

执行以实现所述获取目标服务器对应的附属服务器时，还包括：

解析所述模型训练请求，得到对应目标训练模型的训练数量；

确定所述训练数量对应的训练数量级；

根据所述训练数量级确定附属服务器的数量。

19. 根据权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被处理器执行以实现所述计算所述附属服务器的最优训练数量时，包括：

获取对应每一附属服务器的数据处理能力；

获取所述模型训练请求携带的对应目标训练模型的训练数量；

根据所述数据处理能力与所述训练数量计算每一所述附属服务器的最优训练数量。

20. 根据权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被处理器执行以实现所述基于所述最优训练数量分配所述模型训练请求时，包括：

获取所述附属服务器的数量与每一所述附属服务器的最优训练数量；

基于所述附属服务器的数量与所述最优训练数量拆分所述模型训练请求，得到若干个附属模型训练请求；

将所述附属模型训练请求分配至对应的附属服务器。

1/3

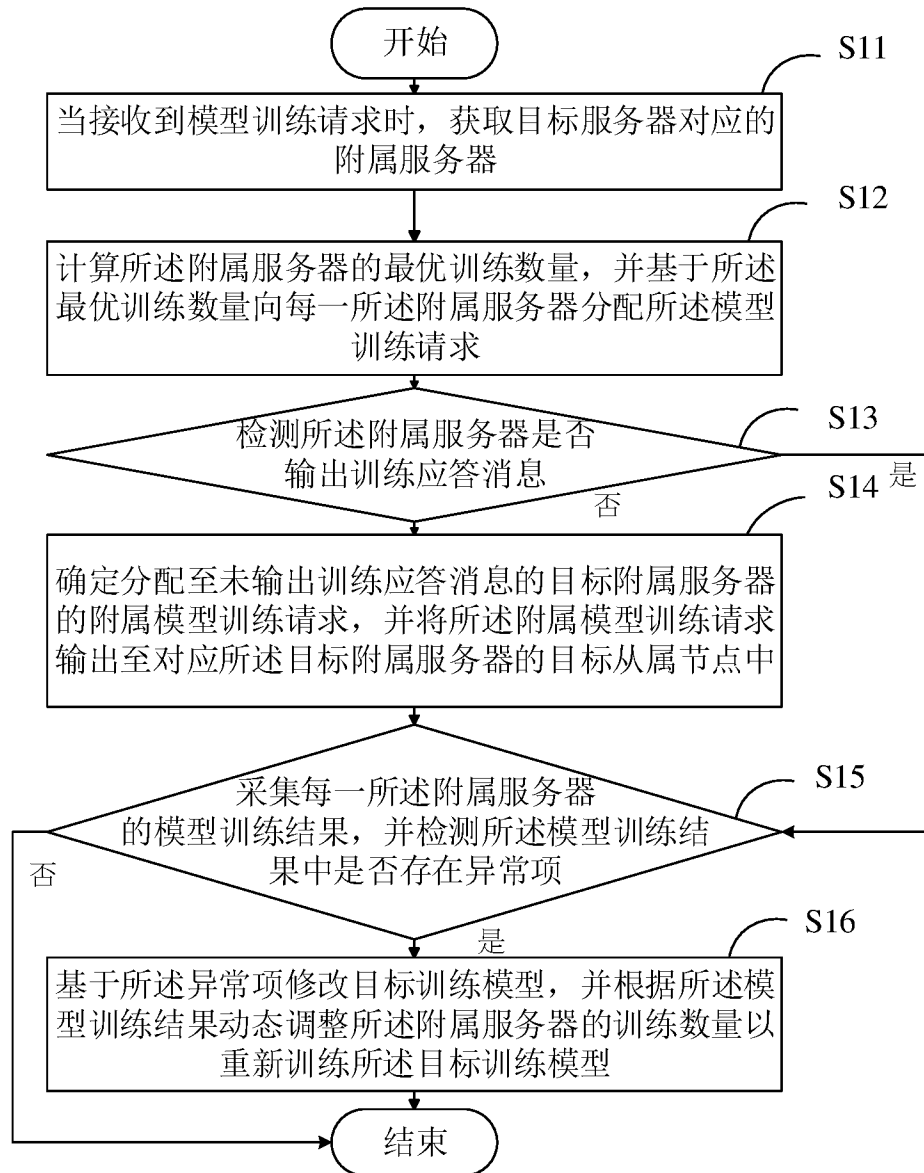


图 1

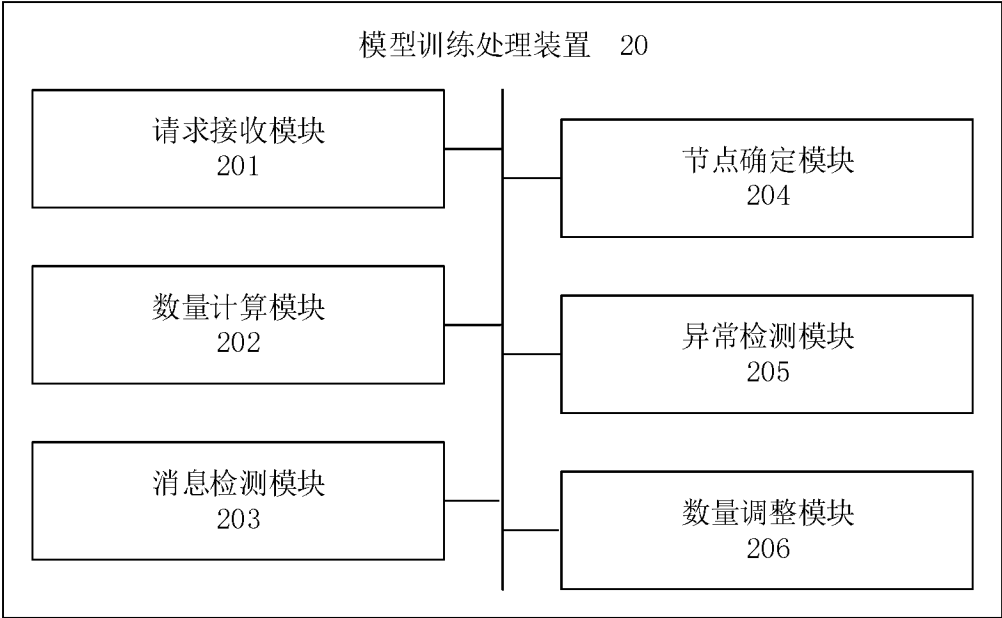


图 2

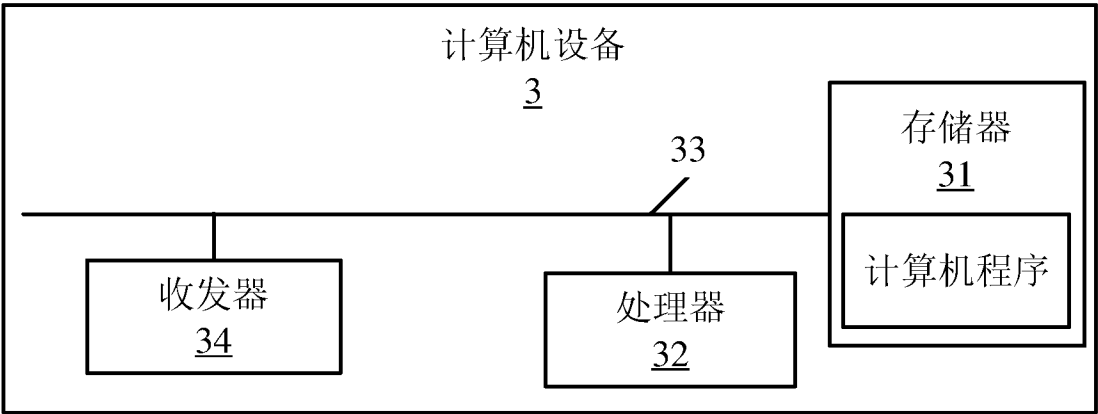


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/125440

A. CLASSIFICATION OF SUBJECT MATTER

G06N 20/00(2019.01)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;G06K

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)

EPTXT; USTXT; DWPI; SIPOABS; WOTXT; CNABS; CNTXT; CNKI: 模型训练, 节点, 分布式, 服务器集群, 深度神经网络, 应答, 异常, model training, node, distributed, server cluster, deep neural network, answer, abnormal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112598135 A (SHENZHEN ONECONNECT TECHNOLOGY CO., LTD.) 02 April 2021 (2021-04-02) description paragraphs 5-111	1-20
A	CN 109543726 A (UNION MOBILE PAY CO., LTD.) 29 March 2019 (2019-03-29) entire document	1-20
A	CN 111753997 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 09 October 2020 (2020-10-09) entire document	1-20
A	CN 111917579 A (UNISOUND INTELLIGENT TECHNOLOGY CO., LTD. et al.) 10 November 2020 (2020-11-10) entire document	1-20
A	US 10152676 B1 (AMAZON TECH INC.) 11 December 2018 (2018-12-11) entire document	1-20



Further documents are listed in the continuation of Box C.



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

23 November 2021

Date of mailing of the international search report

26 January 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/125440

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	112598135	A	02 April 2021	None	
CN	109543726	A	29 March 2019	None	
CN	111753997	A	09 October 2020	CN 111753997 B	27 August 2021
CN	111917579	A	10 November 2020	None	
US	10152676	B1	11 December 2018	None	

国际检索报告

国际申请号

PCT/CN2021/125440

A. 主题的分类

G06N 20/00 (2019.01) i

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

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06N; G06K

包含在检索领域中的除最低限度文献以外的检索文献

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))

EPTXT; USTXT; DWPI; SIPOABS; WOTXT; CNABS; CNTXT; CNKI: 模型训练, 节点, 分布式, 服务器集群, 深度神经网络, 应答, 异常, model training, node, distributed, server cluster, deep neural network, answer, abnormal

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 112598135 A (深圳壹账通智能科技有限公司) 2021年4月2日 (2021 - 04 - 02) 说明书第5-111段	1-20
A	CN 109543726 A (联动优势科技有限公司) 2019年3月29日 (2019 - 03 - 29) 全文	1-20
A	CN 111753997 A (北京百度网讯科技有限公司) 2020年10月9日 (2020 - 10 - 09) 全文	1-20
A	CN 111917579 A (云知声智能科技股份有限公司等) 2020年11月10日 (2020 - 11 - 10) 全文	1-20
A	US 10152676 B1 (AMAZON TECH INC) 2018年12月11日 (2018 - 12 - 11) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“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)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“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

“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

“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

“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

“&” 同族专利的文件

国际检索实际完成的日期

2021年11月23日

国际检索报告邮寄日期

2022年1月26日

ISA/CN的名称和邮寄地址

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

传真号 (86-10)62019451

受权官员

王艳坤

电话号码 86-(10)-62412330

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/125440

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	112598135	A	2021年4月2日	无	
CN	109543726	A	2019年3月29日	无	
CN	111753997	A	2020年10月9日	CN 111753997 B	2021年8月27日
CN	111917579	A	2020年11月10日	无	
US	10152676	B1	2018年12月11日	无	