

获取人工智能模型的方法、装置、设备及存储介质

本申请要求于 2020 年 03 月 31 日提交的申请号为 202010246686.8、发明名称为“获取人工智能模型的方法、装置、设备及存储介质”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

本申请涉及人工智能技术领域，特别涉及一种获取人工智能模型的方法、装置、设备及存储介质。

背景技术

随着人工智能技术的发展，各种人工智能（artificial intelligence，AI）模型越来越多，获取人工智能模型的方式也越来越多。通过联合学习（federated learning，FL）的方式获取人工智能模型便是其中的一种。

联合学习是一种新兴的人工智能基础技术，原本用于解决安卓手机终端用户在本地更新模型的问题，目标是保护终端数据和个人数据隐私、保证合法合规的前提下，在多参与方或多计算结点之间开展高效率的机器学习。目前，联合学习被拓展为在数据不共享的情况下共同构建 AI 模型，提升 AI 模型效果。

发明内容

本申请实施例提供了一种获取人工智能模型的方法、装置、设备及存储介质，以解决相关技术提供的问题，技术方案如下：

第一方面，提供了一种获取人工智能模型的方法，以客户端侧执行该方法为例，该方法包括：客户端接收服务端发送的第一人工智能AI模型，第一AI模型包括多个神经元；客户端在多个神经元中确定参与本轮训练的目标神经元，本轮训练为非首轮训练，目标神经元的数量小于多个神经元的总数量；客户端基于本地数据对目标神经元进行训练；客户端向服务端返回目标神经元对应的参数数据，目标神经元对应的参数数据用于服务端获取收敛的目标AI模型。

由于本申请实施例提供的方法并不需要对模型中的所有神经元均进行训练，而是按照活跃条件来选择部分神经元进行训练，从而能够降低功耗，加快训练速度。此外，又由于客户端向服务端仅传输训练的神经元对应的参数，因而能够节省通信带宽。

在示例性实施例中，目标神经元为活跃度满足条件的神经元。示例性地，目标神经元的活跃度大于活跃度阈值。

在示例性实施例中，多个神经元中的各个神经元具有对应的冷活标志位，冷活标志位的值为第一值时，用于指示神经元参与本轮训练；所述在所述多个神经元中确定参与本轮训练的目标神经元，包括：将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元。

通过冷活标志位来指示活跃条件，从而通过该冷活标志位即可直接确定目标神经元，使

得确定目标神经元的速度较快，进一步节省了训练时间，加快了训练速度。

在示例性实施例中，多个神经元中的各个神经元具有对应的冷冻期，所述冷冻期用于指示神经元不参与训练的周期；

任一神经元对应的冷活标志位的值，基于所述任一神经元对应的冷冻期确定。

在示例性实施例中，所述将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元之前，还包括：对于所述多个神经元中的任一神经元，获取所述任一神经元在上轮训练的活跃度，所述活跃度用于指示神经元被本地数据影响的程度；

根据所述任一神经元在上轮训练的活跃度更新所述任一神经元的冷冻期，根据所述任一神经元的冷冻期更新所述任一神经元的冷活标志位的值。

在示例性实施例中，所述多个神经元中的各个神经元还具有对应的冷冻期计数器，所述冷冻期计数器用于指示所述冷冻期是否结束；

所述获取所述任一神经元在上轮训练的活跃度，包括：响应于所述任一神经元的冷冻期计数器的值指示所述任一神经元的冷冻期结束，获取所述任一神经元在上轮训练的活跃度。

在示例性实施例中，所述在所述多个神经元中确定参与本轮训练的目标神经元，包括：

对于所述多个神经元中的任一神经元，获取所述任一神经元在上轮训练的活跃度，所述活跃度用于指示神经元被本地数据影响的程度；

响应于所述任一神经元在上轮训练的活跃度大于活跃度阈值，将所述任一神经元确定为参与本轮训练的目标神经元。

在示例性实施例中，所述获取所述任一神经元在上轮训练的活跃度，包括：

获取所述任一神经元在上轮训练前得到的参数的第一平均值，以及所述任一神经元在上轮训练后得到的参数的第二平均值；

获取所述第一平均值及所述第二平均值之间的差值；

根据所述差值的绝对值与所述第一平均值的绝对值确定所述任一神经元在上轮训练的活跃度。

在示例性实施例中，所述接收所述服务端发送的第一人工智能AI模型之前，还包括：

接收所述服务端发送的第一人工智能AI模型，将所述多个神经元中的全部神经元确定为目标神经元，基于本地数据对所述目标神经元进行训练。

在示例性实施例中，所述向所述服务端返回所述目标神经元对应的参数数据，还包括：

向所述服务器发送各个神经元的冷活标志位，或者，向所述服务器发送值发生变化的冷活标志位。

在示例性实施例中，所述向所述服务端返回所述目标神经元对应的参数数据，包括：仅向所述服务端返回所述目标神经元对应的参数数据。

提供了一种用于获取人工智能模型的方法，所述方法包括：

服务端获取待训练的第一人工智能AI模型，所述第一AI模型包括多个神经元；

所述服务端向多个客户端发送所述第一AI模型；

所述服务端接收所述多个客户端中的各个客户端返回的目标神经元对应的参数数据，任一客户端返回的目标神经元对应的参数数据由所述任一客户端对所述第一AI模型中的目标神经元进行训练得到，所述目标神经元的数量小于所述多个神经元的总数量；

所述服务端根据所述各个客户端返回的目标神经元对应的参数数据，还原所述各个客户端对应的第二AI模型，根据所述各个客户端对应的第二AI模型获取收敛的目标AI模型。

在示例性实施例中，所述根据所述各个客户端返回的目标神经元对应的参数数据，还原所述各个客户端对应的第二AI模型，包括：

对于任一客户端，根据所述任一客户端返回的目标神经元对应的参数数据，更新所述第一AI模型中的所述目标神经元对应的参数，得到所述任一客户端对应的第二AI模型。

在示例性实施例中，所述根据所述各个客户端对应的第二AI模型获取收敛的目标AI模型，包括：

对所述各个客户端对应的第二AI模型进行联合平均，得到第三AI模型；

响应于所述第三AI模型收敛，将所述第三AI模型作为目标AI模型；

响应于所述第三AI模型未收敛，向所述多个客户端发送所述第三AI模型，采用获取所述第三AI模型的方式继续获取新的AI模型，依此循环，直至得到收敛的目标AI模型。

在示例性实施例中，所述根据所述各个客户端返回的目标神经元对应的参数数据，还原所述各个客户端对应的第二AI模型之前，还包括：

接收所述多个客户端返回的各个神经元的冷活标志位或者值发生变化的冷活标志位；

通过所述各个神经元的冷活标志位或者值发生变化的冷活标志位确定所述多个神经元中的目标神经元。

提供了一种用于获取人工智能模型的装置，所述装置包括：

通信单元，用于接收所述服务端发送的第一人工智能AI模型，所述第一AI模型包括多个神经元；

处理单元，用于在所述多个神经元中确定参与本轮训练的目标神经元，本轮训练为非首轮训练，所述目标神经元的数量小于所述多个神经元的总数量；

所述处理单元，还用于基于本地数据对所述目标神经元进行训练；

所述通信单元，还用于向所述服务端返回所述目标神经元对应的参数数据，所述目标神经元对应的参数数据用于所述服务端获取收敛的目标AI模型。

在示例性实施例中，所述多个神经元中的各个神经元具有对应的冷活标志位，冷活标志位的值为第一值时，用于指示神经元参与本轮训练；

所述处理单元，用于将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元。

在示例性实施例中，所述多个神经元中的各个神经元还具有对应的冷冻期，所述冷冻期用于指示神经元不参与训练的周期；

任一神经元对应的冷活标志位的值，基于所述任一神经元对应的冷冻期确定。

在示例性实施例中，所述处理单元，还用于对于所述多个神经元中的任一神经元，获取所述任一神经元在上轮训练的活跃度，所述活跃度用于指示神经元被本地数据影响的程度；根据所述任一神经元在上轮训练的活跃度更新所述任一神经元的冷冻期，根据所述任一神经元的冷冻期更新所述任一神经元的冷活标志位的值。

在示例性实施例中，所述多个神经元中的各个神经元还具有对应的冷冻期计数器，所述冷冻期计数器用于指示所述冷冻期是否结束；

所述处理单元，用于响应于所述任一神经元的冷冻期计数器的值指示所述任一神经元的

冷冻期结束，获取所述任一神经元在上轮训练的活跃度。

在示例性实施例中，所述处理单元，用于响应于本轮训练为非首轮训练，对于所述多个神经元中的任一神经元，获取所述任一神经元在上轮训练的活跃度，所述活跃度用于指示神经元被本地数据影响的程度；响应于所述任一神经元在上轮训练的活跃度大于活跃度阈值，将所述任一神经元确定为参与本轮训练的目标神经元。

在示例性实施例中，所述处理单元，用于获取所述任一神经元在上轮训练前得到的参数的第一平均值，以及所述任一神经元在上轮训练后得到的参数的第二平均值；获取所述第一平均值及所述第二平均值之间的差值；根据所述差值的绝对值与所述第一平均值的绝对值确定所述任一神经元在上轮训练的活跃度。

在示例性实施例中，所述处理单元，还用于接收所述服务端发送的第一人工智能AI模型，将所述多个神经元中的全部神经元确定为目标神经元，基于本地数据对所述目标神经元进行训练。

在示例性实施例中，所述通信单元，还用于向所述服务器发送各个神经元的冷活标志位，或者，向所述服务器发送值发生变化的冷活标志位。

在示例性实施例中，所述通信单元，用于仅向所述服务端返回所述目标神经元对应的参数数据。

还提供了一种用于获取人工智能模型的装置，所述装置包括：

处理单元，用于获取待训练的第一人工智能AI模型，所述第一AI模型包括多个神经元；

通信单元，用于向多个客户端发送所述第一AI模型；

所述通信单元，还用于接收所述多个客户端中的各个客户端返回的目标神经元对应的参数数据，任一客户端返回的目标神经元对应的参数数据由所述任一客户端对所述第一AI模型中的目标神经元进行训练得到，所述目标神经元的数量小于所述多个神经元的总数量；

所述处理单元，还用于根据所述各个客户端返回的目标神经元对应的参数数据，还原所述各个客户端对应的第二AI模型，根据所述各个客户端对应的第二AI模型获取收敛的目标AI模型。

在示例性实施例中，所述处理单元，用于对于任一客户端，根据所述任一客户端返回的目标神经元对应的参数数据，更新所述第一AI模型中的所述目标神经元对应的参数，将所述第一AI模型中的非目标神经元对应的参数进行补齐，得到所述任一客户端对应的第二AI模型。

在示例性实施例中，所述处理单元，用于对所述各个客户端对应的第二AI模型进行联合平均，得到第三AI模型；响应于所述第三AI模型收敛，将所述第三AI模型作为目标AI模型；响应于所述第三AI模型未收敛，向所述多个客户端发送所述第三AI模型，采用获取所述第三AI模型的方式继续获取新的AI模型，依此循环，直至得到收敛的目标AI模型。

在示例性实施例中，所述通信单元，还用于接收所述多个客户端返回的各个神经元的冷活标志位或者值发生变化的冷活标志位；

所述处理单元，用于通过所述各个神经元的冷活标志位或者值发生变化的冷活标志位确定所述多个神经元中的目标神经元。

还提供一种获取人工智能模型的设备，该设备包括：存储器及处理器，所述存储器中存

储有至少一条指令，所述至少一条指令由所述处理器加载并执行，以实现上述第一方面或第一方面任一所述的获取人工智能模型的方法。

还提供一种获取人工智能模型的设备，该设备包括：存储器及处理器，所述存储器中存储有至少一条指令，所述至少一条指令由所述处理器加载并执行，以实现上述第二方面或第二方面任一所述的获取人工智能模型的方法。

还提供了一种计算机可读存储介质，所述存储介质中存储有至少一条指令，所述指令由处理器加载并执行以实现如上任一所述的获取人工智能模型的方法。

提供了另一种通信装置，该装置包括：收发器、存储器和处理器。其中，该收发器、该存储器和该处理器通过内部连接通路互相通信，该存储器用于存储指令，该处理器用于执行该存储器存储的指令，以控制收发器接收信号，并控制收发器发送信号，并且当该处理器执行该存储器存储的指令时，使得该处理器执行第一方面或第一方面的任一种可能的实施方式中的方法。

提供了另一种通信装置，该装置包括：收发器、存储器和处理器。其中，该收发器、该存储器和该处理器通过内部连接通路互相通信，该存储器用于存储指令，该处理器用于执行该存储器存储的指令，以控制收发器接收信号，并控制收发器发送信号，并且当该处理器执行该存储器存储的指令时，使得该处理器执行第二方面或第二方面的任一种可能的实施方式中的方法。

作为一种示例性实施例，所述处理器为一个或多个，所述存储器为一个或多个。

作为一种示例性实施例，所述存储器可以与所述处理器集成在一起，或者所述存储器与处理器分离设置。

在具体实现过程中，存储器可以为非瞬时性（non-transitory）存储器，例如只读存储器（read only memory, ROM），其可以与处理器集成在同一块芯片上，也可以分别设置在不同的芯片上，本申请实施例对存储器的类型以及存储器与处理器的设置方式不做限定。

提供了一种计算机程序（产品），所述计算机程序（产品）包括：计算机程序代码，当所述计算机程序代码被计算机运行时，使得所述计算机执行上述各方面中的方法。

提供了一种芯片，包括处理器，用于从存储器中调用并运行所述存储器中存储的指令，使得安装有该芯片的通信设备执行上述各方面中的方法。

提供另一种芯片，包括：输入接口、输出接口、处理器和存储器，所述输入接口、输出接口、所述处理器以及所述存储器之间通过内部连接通路相连，所述处理器用于执行所述存储器中的代码，当所述代码被执行时，所述处理器用于执行上述各方面中的方法。

附图说明

图1为本申请实施例提供的联合学习的网络架构示意图；

图2为本申请实施例提供的联合学习的网络架构示意图；

图3为本申请实施例提供的用于获取人工智能模型的方法流程图；

图4为本申请实施例提供的获取人工智能模型的交互过程示意图；

图5为本申请实施例提供的客户端侧用于获取人工智能模型的过程示意图；

图6为本申请实施例提供的用于获取人工智能模型的装置结构示意图；

图7为本申请实施例提供的用于获取人工智能模型的装置结构示意图；

图8为本申请实施例提供的用于获取人工智能模型的设备结构示意图。

具体实施方式

本申请的实施方式部分使用的术语仅用于对本申请的实施例进行解释，而非旨在限定本申请。

随着人工智能技术的发展，各种人工智能模型越来越多，获取人工智能模型的方式也越来越多，通过联合学习的方式获取AI模型便是其中的一种。联合学习是一种新兴的人工智能基础技术，原本用于解决安卓手机终端用户在本地更新模型的问题，目标是保护终端数据和个人数据隐私、保证合法合规的前提下，在多参与方或多计算结点之间开展高效率的机器学习。目前，联合学习被拓展为在数据不共享的情况下共同构建AI模型，提升AI模型效果。

以如图1所示的联合学习的网络架构为例，该网络架构分为两层：下层为多个客户端，上层为一个服务端。其中，客户端一般指一个设备，该设备会流入诸多数据。如图1所示，以该网络架构中包括的多个客户端分别为客户端1、客户端2和客户端3为例，一般情况下，各客户端数据很难采集或共享，因而在客户端1数据下训练的AI模型仅对客户端1有效，在客户端2数据下训练的AI模型仅对客户端2有效，在客户端3数据下训练的AI模型仅对客户端3有效。为了能够利用所有客户端的数据，通过联合学习来获取AI模型的方式应运而生。

在通过联合学习来获取AI模型的过程中，各客户端利用本设备的数据（各客户端数据可以不同），在一个初始中心模型（一般相同）上，启动本地训练。各客户端本地训练多批次后，将本地模型上传到服务端；服务端对各本地模型进行汇聚，并进行联合更新，形成更新的中心模型。如果更新的中心模型未收敛，则服务端将该更新的中心模型继续下发给各客户端。各客户端利用本设备的数据，在更新的中心模型上继续本地训练，再上传服务端；如此循环，直至得到收敛的中心模型，联合训练完毕。

通过上述联合学习的方式进行训练得到的中心模型，对所有客户端的数据都有优秀的判别能力，极大程度上满足了采集不同客户端的数据，就可以共同构建一个AI模型的需求。

然而，由于上述通过联合学习训练AI模型的过程中，每个客户端均要对服务端下发的完整模型进行完整训练，训练时间较长，且客户端与服务端之间传输的也是完整的模型，导致通信开销较大。

对此，本申请实施例提出了一种获取AI模型的方法，该方法在客户端进行本地训练时，仅训练能够被本地数据影响而活跃的神经元，冷冻不活跃的神经元。其中，冷冻不活跃的神经元是指该神经元不参与训练。由于每次客户端仅上传活跃的目标神经元的参数数据，以将重要信息传给服务端，那么即使是小规模的客户端，服务端也不会丢失融合的重要信息。另外，由于训练时仅训练非冷冻的目标神经元，客户端的训练时间得以加快、设备功耗得以降低。

例如，本申请实施例提供的方法可应用于图2所示的网络架构中。如图2所示，该网络架构包括客户端1和客户端2这两个客户端，该客户端1和客户端2可以是任何支持本地训练的设备，包括但不限于：手机、平板、电脑、交换机、光线路终端(optical line terminal, OLT)、光网络终端(optical network terminal, ONT)、路由器等。服务端可指任何支持多模型汇聚、联合的设备或云平台，包括但不限于：服务器、公有云等。

如图2所示，各客户端以本地数据来进行本轮的模型训练，仅对本轮参与训练的目标神经

元进行训练，训练结束后，将目标神经元对应的参数数据上传至服务端。例如，将本地模型上传到服务端，或者仅上传本轮参与训练的目标神经元对应的参数，或者上传各个神经元对应的参数（包括目标神经元对应的参数和非目标神经元对应的参数），或者仅上传目标神经元对应的参数差值。服务端汇聚各个客户端返回的参与训练的目标神经元对应的参数数据之后，还原出各个客户端的AI模型，据此得到更新的中心模型。如果该更新的中心模型未收敛，再将更新的中心模型下发给各客户端。各客户端再开始下一轮的训练，直至得到收敛的中心模型为止。

接下来，以上述图 2 所示的网络架构为例，对本申请实施例提供的获取人工智能模型的方法进行说明。如图 3 所示，该方法可通过服务端和客户端的交互来实现 AI 模型的获取，本申请实施例提供的方法可包括如下几个过程。

301，服务端获取待训练的第一AI模型，第一AI模型包括多个神经元。

其中，待训练的第一AI模型可以是未经过训练的初始AI模型，也可以是经过一轮或多轮训练，但未收敛，仍然需要再训练的AI模型。无论是初始AI模型，还是未收敛需要再训练的AI模型，该第一AI模型均包括多个神经元。神经元是构成AI模型的基本单元，其主要是模拟生物神经元的结构和特性，接受一组输入信号并产出输出。

本申请实施例不对第一AI模型包括的神经元的数量进行限定，且本申请实施例不对待训练的第一AI模型的种类进行限定，该第一AI模型可以是AI模型中的任一种类型。例如，服务器在获取该待训练的第一AI模型时，可基于应用场景来确定获取哪种类型的第一AI模型。

302，服务端向多个客户端发送第一 AI 模型。

在本申请实施例提供的方法所应用的实施环境中，如图 2 所示，服务端与多个客户端进行通信连接，该多个客户端可进行联合学习，以获取最终收敛的目标 AI 模型。因此，服务端获取第一 AI 模型，该服务端可将该第一 AI 模型发送给多个客户端，由多个客户端分别执行对该第一 AI 模型的训练过程。

在示例性实施例中，服务端在向客户端发送第一 AI 模型时，还可以向客户端发送训练的批次数，以使各个客户端能够确定对第一 AI 模型进行本轮训练的次数。如果需要多轮训练，服务端可以每次向客户端发送第一 AI 模型时，均发送对应的批次数，不同轮的批次数可以不同，也可以相同。如果每轮训练的批次数相同，服务端也可以仅在首轮训练时向客户端下发批次数，后续不再重复发送。另外，除了服务端向客户端发送训练的批次数的方式外，服务端也可以不下发，而是按照提前约定好的批次数进行每轮训练。

303，客户端接收服务端发送的第一AI模型。

该客户端可以是与服务端进行连接的多个客户端中的任一客户端。

需要说明的是，本申请实施例中仅以与服务端连接的多个客户端中的一个客户端为例，对客户端的操作过程进行说明。本申请实施例中涉及的客户端的操作过程，对于服务端连接的多个客户端均适用，不再对每个客户端的操作过程一一进行说明。

304，客户端在多个神经元中确定参与本轮训练的目标神经元，目标神经元的数量小于多个神经元的总数量，本轮训练为非首轮训练。

由于本地数据对不同神经元的影响程度不同，有些神经元经过本地数据训练后，可能参数变化不大，因而在本申请实施例提供的方法中，客户端接收到待训练的第一AI模型后，可

以无需对该第一AI模型中的所有神经元均进行训练，而是仅训练活跃度满足一定条件的目标神经元。活跃度满足一定条件的目标神经元可以是能够通过本地数据影响而活跃的神经元。关于客户端在多个神经元中确定参与本轮训练的目标神经元的方式，本申请实施例不进行限定。示例性地，目标神经元是活跃度满足条件的神经元，例如，活跃度大于活跃度阈值。

针对首轮训练的情况，由于未对第一AI模型进行训练，因而还无法判断哪些神经元活跃，哪些神经元不活跃，因此，在多个神经元中确定参与本轮训练的目标神经元，包括：接收所述服务端发送的第一人工智能AI模型，将所述多个神经元中的全部神经元确定为目标神经元，基于本地数据对所述目标神经元进行训练。

针对本轮训练为非首轮训练的情况，仅以如下两种方式进行举例说明。

方式一：基于神经元的活跃度，在多个神经元中确定参与本轮训练的目标神经元。

对于本轮训练为非首轮训练的情况，由于上轮训练已经采用本地数据对神经元进行过训练，可以确定哪些神经元活跃，哪些神经元不活跃。因此，针对该种本轮训练为非首轮训练的情况，在多个神经元中确定参与本轮训练的目标神经元，包括但不限于：对于多个神经元中的任一神经元，获取任一神经元在上轮训练的活跃度，活跃度用于指示神经元被本地数据影响的程度；响应于任一神经元在上轮训练的活跃度大于活跃度阈值，将任一神经元作为参与本轮训练的目标神经元。

关于获取任一神经元在上轮训练的活跃度的方式，本申请实施例不进行限定，示例性地，获取任一神经元在上轮训练的活跃度，包括：获取任一神经元在上轮训练前得到的参数的第一平均值，以及任一神经元在上轮训练后得到的参数的第二平均值；获取第一平均值及第二平均值之间的差值；根据差值的绝对值与第一平均值的绝对值确定作为任一神经元在上轮训练的活跃度。

例如，以任一神经元在上轮训练前得到的参数的第一平均值为AC，该任一神经元在上轮训练后得到的参数的第二平均值为AE为例，则将差值的绝对值与第一平均值的绝对值的商确定为任一神经元在上轮训练的活跃度。例如，通过如下公式获取该任一神经元在上轮训练的活跃度AR：

$$AR = \frac{|AE-AC|}{|AC|} \text{ 或者, } AR = \frac{|AE-AC|}{|AC|+x}$$

除采用上述获取任一神经元在上轮训练的活跃度的方式外，还可以基于上述方式进行变型，已得到可选的其他获取活跃度的方式。其中，x为防止商为零的极小值，例如，可以是0.0001。本申请实施例不对x的值进行限定。

例如，仍以任一神经元在上轮训练前得到的参数的第一平均值为AC，该任一神经元在上轮训练后得到的参数的第二平均值为AE为例，通过如下公式获取该任一神经元在上轮训练的活跃度AR：

$$AR = \frac{|AE-AC|}{|AC|} * \mu \text{ 或者, } AR = \frac{|AE-AC|}{|AC|+x} * \mu$$

其中， μ 为活跃度系数，可基于经验设置，也可以根据每轮训练的情况进行调整。

又例如，仍以任一神经元在上轮训练前得到的参数的第一平均值为AC，该任一神经元在上轮训练后得到的参数的第二平均值为AE为例，通过如下公式获取该任一神经元在上轮训练的活跃度AR：

$$AR = \frac{|AE-AC|}{|AC|} + \varphi \text{ 或者 } AR = \frac{|AE-AC|}{|AC|+x} + \varphi$$

其中, φ 为活跃度常数, 可基于经验设置, 也可以根据每轮训练的情况进行调整。

无论采用上述哪种方式获取任一神经元的活跃度, 均可以将获取的活跃度与活跃度阈值进行比对, 如果活跃度大于活跃度阈值, 则将该任一神经元作为参与本轮训练的目标神经元。关于活跃度阈值, 本申请实施例不进行限定, 可基于经验设置, 也可以根据每轮训练的情况进行调整。

方式二: 通过查询冷活标志位, 在多个神经元中确定参与本轮训练的目标神经元。

该方式二与上述方式一每轮训练时计算神经元的活跃度, 通过活跃度是否满足阈值来确定目标神经元的不同之处在于, 本申请实施例提供的方法可以在本轮训练时不计算活跃度, 而是通过查询冷活标志位的方式来实现。在示例性实施例中, 多个神经元中的各个神经元具有对应的冷活标志位, 冷活标志位的值为第一值时, 用于指示神经元参与本轮训练。例如, 每轮训练结束后, 可采用上述任一方式计算活跃度, 将活跃度满足阈值的神经元的冷活标志位的值设为第一值。则在多个神经元中确定参与本轮训练的目标神经元, 包括: 将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元。

上述通过冷活标志位来指示是否参与本轮训练的方式, 虽然冷活标志位的值也需要通过活跃度来确定, 但由于是在上轮训练结束后, 也即在本轮训练前计算得到的, 而在本轮训练时, 通过该冷活标志位即可直接确定目标神经元, 使得确定参加本轮训练的目标神经元的速度较快, 进一步节省了训练时间, 加快了训练速度。

进一步地, 关于本轮训练时, 神经元的冷活标志位的值是如何确定的, 本申请实施例不进行限定。在示例性实施例中, 多个神经元中的各个神经元还具有对应的冷冻期, 冷冻期用于指示神经元不参与训练的周期。任一神经元对应的冷活标志位的值, 基于所述任一神经元对应的冷冻期确定。例如, 任一神经元的冷冻期结束, 则该任一神经元的冷活标志位为第一值。如此, 通过设置冷冻期来提供一种探测机制, 以基于冷冻期来指示神经元不参与训练的周期, 基于冷冻周期来确定神经元是否参与训练, 确定冷活标志位的值, 从而防止丢失在训练过程中需要参与训练的神经元对应的参数, 保证了鲁棒性。

则本申请实施例将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元之前, 还包括: 对于多个神经元中的任一神经元, 获取任一神经元在上轮训练的活跃度, 活跃度用于指示神经元被本地数据影响的程度; 根据任一神经元在上轮训练的活跃度更新任一神经元的冷冻期, 根据任一神经元的冷冻期更新任一神经元的冷活标志位的值。

其中, 获取任一神经元在上轮训练的活跃度的方式, 可参考上述方式一中有关活跃度计算方式的介绍, 此处不再赘述。示例性地, 根据任一神经元在上轮训练的活跃度更新任一神经元的冷冻期, 包括但不限于通过如下公式来更新冷冻期:

$$Ti' = \begin{cases} Ti + 1 & ARi < X1 \\ Ti & X1 \leq ARi \leq X2 \\ \lfloor (Ti + 1) / 2 \rfloor & ARi \geq X2 \end{cases}$$

上述更新冷冻期的公式中, Ti' 代表第 i 个神经元更新后的冷冻期, Ti 代表第 i 个神经元更新前的冷冻期, ARi 代表第 i 个神经元在上轮训练的活跃度, i 为 1 到 N 之间的整数, N 为神经元的总数量。 $X1$ 代表活跃度第一阈值, $X2$ 代表活跃度第二阈值, $X2$ 大于 $X1$ 。 $\lfloor (Ti + 1) / 2 \rfloor$ 代表对 $(Ti + 1) / 2$ 进行向下取整。

关于活跃度第一阈值及活跃度第二阈值的大小，本申请实施例不进行限定，可基于经验设置，也可以基于训练情况进行调整。例如， $X1=0.2$ ， $X2=0.4$ ，则上述更新冷冻期的公式为：

$$Ti' = \begin{cases} Ti + 1 & ARi < 0.2 \\ Ti & 0.2 \leq ARi \leq 0.4 \\ \lfloor (Ti + 1) / 2 \rfloor & ARi \geq 0.4 \end{cases}$$

以任一神经元的活跃度为0.3为例，则该神经元更新后的冷冻期 Ti' 与更新前的冷冻期 Ti 的值相同。可以认为该神经元的活跃度没有发生太大的变化，因而可以保持原有冷冻期。以任一神经元的活跃度为0.1为例，则该神经元更新后的冷冻期 $Ti'=Ti+1$ ，可以认为该神经元不够活跃，因而增加冷冻期。以任一神经元的活跃度为0.6为例，则该神经元更新后的冷冻期 $Ti'=\lfloor (Ti + 1) / 2 \rfloor$ ，可以认为该神经元比较活跃，因而减少冷冻期，以尽快使该神经元参与训练，进而保证不丢失重要信息。

在示例性实施例中，根据任一神经元的冷冻期更新任一神经元的冷活标志位的值，包括但不限于对 $Ti' > Y$ 的神经元，令 Fi' 为第二值，例如第二值为1，代表本轮训练可以被冷冻，即不参与本轮训练；对 $Ti'=Y$ 的神经元，令 Fi' 为第一值，例如第一值为0，代表本轮训练不被冷冻，即参与本轮训练。示例性地， Y 可以为1。

在示例性实施例中，在一轮训练结束后，可以不对每个神经元的冷冻期均实施上述更新过程，而仅对冷冻期结束的神经元进行更新。多个神经元中的各个神经元还具有对应的冷冻期计数器，冷冻期计数器用于指示冷冻期是否结束；获取任一神经元在上轮训练的活跃度，包括：响应于任一神经元的冷冻期计数器的值指示任一神经元的冷冻期结束，获取任一神经元在上轮训练的活跃度。

该种实施方式下，通过设置冷冻期计数器，基于冷冻期计数器来指示冷冻期是否结束，并在冷冻期结束后，再获取活跃度，进一步节省了探测冷冻期是否结束的时间，提高了训练速度。示例性地，活跃度越低，则冷冻期的值越大，冷冻期越长。其中，冷冻期计数器可以初始值设为1，训练轮数每增加一轮，则冷冻期计数器的值加1，也就是说，冷冻期的值与冷冻期计数器的值正相关。则冷冻期计数器的值与冷冻期的值相等时，可认为冷冻期结束。除此之外，冷冻期计数器的值还可以与冷冻期的值负相关，例如，活跃度越低，则冷冻期的值越大，冷冻期越长。其中，冷冻期计数器可以初始值设为冷冻期的值，训练轮数每增加一轮，则冷冻期计数器的值减1。则冷冻期计数器的值为零时，可认为冷冻期结束。

需要说明的是，针对本轮训练为首轮训练的情况，由于无法基于上轮训练结果确定神经元在上轮训练的活跃度，也就无法基于上轮训练的活跃度来确定神经元在本轮训练时的冷活标志位的值。对此，本申请实施例提供了首轮训练时的初始化过程。例如，响应于本轮训练为首轮训练，将多个神经元中的各个神经元的冷活标志位的值置为第一值。

也就是说，在本轮训练为首轮训练的情况下，每个神经元的冷活标志位的值均设置为第一值，那么每个神经元均可被确定为是目标神经元，从而参与本轮训练。之后便可基于本轮训练结果来更新神经元的冷活标志位的值，以使下轮训练时，能够根据本轮更新的冷活标志位的值来直接确定目标神经元。关于更新神经元的冷活标志位的值的方式，可参见上面的相关描述，此处暂不赘述。

305，客户端基于本地数据对目标神经元进行训练。

在本申请实施例提供的方法中，客户端基于本地数据对目标神经元进行训练，可以根据本轮训练的批次来进行。例如，本轮训练包括E批次，则客户端对目标神经元进行一批次训练

之后，得到更新后的AI模型，之后继续进行训练，直至训练E批次结束。

其中，本轮训练的批次数可以是服务端发送的，也可以是提前约定好的，本申请实施例不对本轮训练的批次数获取方式以及次数多少进行限定。

306，客户端向服务端返回目标神经元对应的参数数据。

关于客户端向服务端返回目标神经元对应的参数数据的方式，本申请实施例不进行限定，可以是仅返回目标神经元对应的参数，也可以仅返回目标神经元对应的参数差值，从而减少客户端与服务端之间的通信带宽。除此之外，由于该训练后得到的AI模型包括该目标神经元对应的参数，因而本申请实施例提供的方法同样也支持返回训练后得到的AI模型，通过返回训练后得到的AI模型的方式，来实现返回目标神经元对应的参数数据。也可以向服务端返回全部神经元对应的参数或参数差值。其中，任一神经元对应的参数差值可以是本轮训练得到的该任一神经元对应的参数与上轮训练得到的该任一神经元对应的参数之间的差值。

针对通过冷活标志位来指示神经元是否参与本轮训练的方式，在示例性实施例中，客户端向服务端返回目标神经元对应的参数数据，还包括：向服务器发送各个神经元的冷活标志位，或者，向服务器发送值发生变化的冷活标志位。关于发送冷活标志位的方式，本申请实施例不进行限定。如果本轮训练为首轮训练，则客户端可以向服务器发送各个神经元的冷活标志位，对于后续冷活标志位的值未改变的神经元，客户端可以在后续不再发送该神经元的冷活标志位，从而进一步减少通信带宽。示例性地，发送冷活标志位的方式，本申请实施例不进行限定，例如，如果有5个神经元，可以采用5比特数据来指示，一个比特位对应一个神经元的冷活标志位，对于目标神经元，冷活标志位的第一值为1，则客户端返回的冷活标志位可以是10011。第1个比特位、第4个比特位和第5个比特位的值为1，则说明第1个神经元、第4个神经元和第5个神经元为目标神经元。

307，服务端接收多个客户端中的各个客户端返回的目标神经元对应的参数数据。

其中，任一客户端返回的目标神经元对应的参数数据由任一客户端对第一AI模型中的目标神经元进行训练得到，目标神经元的数量小于多个神经元的总数量。

需要说明的是，虽然服务端可以向多个客户端中的各个客户端发送相同的第一AI模型，但由于各个客户端的本地数据不同，而本地数据对神经元的影响程度不同。因此，不同客户端确定出参加本轮训练的目标神经元有可能会存在不一致的情况，且通过本地数据对目标神经元训练之后得到的参数也会存在不同。也就是说，服务端从多个客户端接收到的目标神经元对应的参数数据有可能是不同的。

例如，以图2所示的网络架构为例，如果第一AI模型包括4个神经元，分别为神经元1、神经元2、神经元3和神经元4。客户端1接收到服务端发送的第一AI模型后，基于客户端1的本地数据确定出这4个神经元的活跃度之后，确定神经元1和神经元2的活跃度大于活跃度阈值，也即客户端1确定出的目标神经元为神经元1和神经元2，则向服务端返回的参数数据也是神经元1的参数数据和神经元2的参数数据。而对于客户端2，接收到服务端发送的第一AI模型后，基于客户端2的本地数据确定出这4个神经元的活跃度之后，确定神经元1、神经元2和神经元3的活跃度均大于活跃度阈值，也即客户端2确定出的目标神经元为神经元1、神经元2和神经元3，则向服务端返回的参数数据也是神经元1的参数数据、神经元2的参数数据和神经元3的参数数据。

由此可见，客户端1确定的目标神经元的数量与客户端2确定的目标神经元的数量不同，

且即使客户端1和客户端2都将神经元1和神经元2确定为目标神经元,但由于客户端1和客户端2的本地数据不同,因此,针对同一目标神经元1和神经元2,客户端1对这两个神经元进行训练得到的参数数据与客户端2对这两个神经元进行训练得到的参数数据也可能会不同。

308, 服务端根据各个客户端返回的目标神经元对应的参数数据,还原各个客户端对应的第二 AI 模型,根据各个客户端对应的第二 AI 模型获取收敛的目标 AI 模型。

如上,由于各个客户端的本地数据可能存在差异,各个客户端向服务端返回的目标神经元对应的参数数据也可能不同,因此,服务端根据各个客户端返回的目标神经元对应的参数数据,分别还原各个客户端对应的第二 AI 模型。

在示例性实施例中,服务端根据各个客户端返回的目标神经元对应的参数数据,还原各个客户端对应的第二 AI 模型,包括:对于任一客户端,根据任一客户端返回的目标神经元对应的参数数据,更新第一 AI 模型中的目标神经元对应的参数,得到任一客户端对应的第二 AI 模型。例如,第二 AI 模型中的非目标神经元为第一 AI 模型中的非目标神经元对应的参数。

例如,以第一 AI 模型包括神经元 1 和神经元 2 为例,客户端 2 返回神经元 1 的参数,未返回神经元 2 的参数。则神经元 1 为目标神经元,神经元 2 为非目标神经元,将客户端返回的该神经元 1 对应的参数更新第一 AI 模型中该神经元 1 对应的参数,对第一 AI 模型中的神经元 2 对应的参数进行补齐,得到该客户端 2 对应的第二 AI 模型。

又例如,仍以第一 AI 模型包括神经元 1 和神经元 2 为例,客户端 1 返回神经元 2 对应的参数,未返回神经元 1 对应的参数。则神经元 2 为目标神经元,神经元 1 为非目标神经元,将客户端返回的该神经元 2 对应的参数更新第一 AI 模型中该神经元 2 对应的参数,对第一 AI 模型中的神经元 1 对应的参数进行补齐,得到该客户端 1 对应的第二 AI 模型。

其中,对非目标神经元对应的参数进行补齐的方式,本申请实施例不进行限定。例如,可将上轮更新的 AI 模型中同一非目标神经元对应的参数,替换该第一 AI 模型中该非目标神经元对应的参数。

需要说明的是,如果客户端返回的是目标神经元对应的参数差值,则可以将第一 AI 模型中该目标神经元对应的参数的值与参数差值综合得到该目标神经元对应的参数。

还原得到各个客户端对应的第二 AI 模型时,为了能够区分哪些神经元是目标神经元,本申请实施例提供的方法在根据各个客户端返回的目标神经元对应的参数,还原各个客户端对应的第二 AI 模型之前,还包括:接收多个客户端返回的各个神经元的冷活标志位或者值发生变化的冷活标志位;通过各个神经元的冷活标志位或者值发生变化的冷活标志位确定多个神经元中的目标神经元。

在示例性实施例中,根据各个客户端对应的第二 AI 模型获取收敛的目标 AI 模型,包括:对各个客户端对应的第二 AI 模型进行联合平均,得到第三 AI 模型;响应于第三 AI 模型收敛,将第三 AI 模型作为目标 AI 模型;响应于第三 AI 模型未收敛,向多个客户端发送第三 AI 模型,采用获取第三 AI 模型的方式继续获取新的 AI 模型,依此循环,直至得到收敛的目标 AI 模型。

示例性地,对各个客户端对应的第二 AI 模型进行联合平均,得到第三 AI 模型的方式,本申请实施例不进行限定。由于一个神经元可能有一个或多个参数,因而可将各个第二 AI 模型的同一神经元对应的同一参数进行平均,将该平均值作为第三 AI 模型中该同一神经元对应的同一参数的参数值。例如,对各个客户端对应的第二 AI 模型中的同一神经元,可取该同

一神经元对应的参数的平均值作为目标参数值,以该目标参数值替换第一 AI 模型中的该神经元对应的参数值,从而得到第三 AI 模型。例如,以第一 AI 模型包括神经元 1 和神经元 2 为例,任一客户端对应的第二 AI 模型是基于第一 AI 模型进行还原得到的。因此,该第二 AI 模型也包括神经元 1 和神经元 2。则将所有第二 AI 模型的神经元 1 对应的第一参数的值取平均值之后,替换第一 AI 模型中的神经元 1 对应的第一参数的值,将所有第二 AI 模型的神经元 2 对应的第二参数的值取平均值之后,替换第一 AI 模型中的神经元 2 对应的第二参数的值,得到第三 AI 模型。

除上述还原方式外,还可以进一步变型得到新的还原方式。例如,对各个客户端对应的第二 AI 模型中的同一神经元,可取该同一神经元对应的参数的平均值,将该平均值与参考系数的乘积作为目标参数值,以该目标参数值替换第一 AI 模型中的该神经元对应的参数值,从而得到第三 AI 模型。仍以第一 AI 模型和第二 AI 模型均包括神经元 1 和神经元 2 为例,则将所有第二 AI 模型的神经元 1 对应的第一参数的值取平均得到第一平均值之后,将第一平均值与参考系数的乘积替换第一 AI 模型中的神经元 1 对应的第一参数的值,将所有第二 AI 模型的神经元 2 对应的第二参数取平均得到第二平均值之后,将第二平均值与参考系数的乘积替换第一 AI 模型中的神经元 2 对应的第二参数,得到第三 AI 模型。其中,参考系数可以基于经验设置,也可以根据应用场景进行调整,本申请实施例不进行限定。

无论采用上述哪种方式进行联合平均得到第三 AI 模型,得到第三 AI 模型之后,判断该第三 AI 模型是否收敛,响应于第三 AI 模型收敛,将第三 AI 模型作为目标 AI 模型。响应于第三 AI 模型未收敛,向多个客户端发送第三 AI 模型,采用获取第三 AI 模型的方式继续获取新的 AI 模型,依此循环,直至得到收敛的目标 AI 模型。

关于第三 AI 模型是否收敛的方式,本申请实施例不进行限定,可以是训练轮数达到训练阈值,也可以是第三 AI 模型的性能达到一定要求。

综上所述,本申请实施例提供的方法,由于并不需要对模型中的所有神经元均进行训练,而是按照活跃条件来选择部分神经元进行训练,从而能够降低功耗,加快训练速度。此外,又由于客户端向服务端仅传输训练的神经元对应的参数,因而能够节省通信带宽。

另外,通过冷活标志位来指示活跃条件,从而通过该冷活标志位即可直接确定目标神经元,使得确定目标神经元的速度较快,进一步节省了训练时间,加快了训练速度。通过设置冷冻期,基于冷冻期来指示神经元不参与训练的周期,基于冷冻周期来确定神经元是否参与训练,从而防止丢失在训练过程中需要参与训练的神经元对应的参数,保证了鲁棒性。通过设置冷冻期计数器,基于冷冻期计数器来指示冷冻期是否结束,并在冷冻期结束后,再获取活跃度,进一步节省了探测冷冻期是否结束的时间,提高了训练速度。

本申请实施例提供的获取人工智能模型的方法整体流程可以图 4 所示的交互过程为例,接下来以服务端与多个客户端中的目标客户端的交互过程为例,对本申请实施例提供的获取人工智能模型的方法进行说明。如图 4 所示,该方法包括但不限于如下几个过程:

401,服务端下发中心模型给多个客户端,目标客户端接收该中心模型,该中心模型有 N 个神经元。

402,客户端开始本地训练,若本回是第一回,则初始化冷冻期 $T_i=1$, $i \in \{1,2,\dots,N\}$,代表第 i 个神经元在本回合进行冷冻探测;初始化活冷标志位 $F_i=0$, $i \in \{1,2,\dots,N\}$,代表第 i 个

神经元没有冷冻（若为 1，则代表冷冻）；初始化冷冻期间计数器 $T_{ci}=1$, $i \in \{1,2,\dots,N\}$ （每和服务端交互一回, $T_{ci}+1$ ）；开始本回本地基于神经元冷冻法的训练，更新 T_i 、 F_i 、 T_{ci} 值。

403，各客户端上传参与本轮训练的目标神经元的参数数据至服务端。

404，若活冷标志位 F_i 发生改变，则目标客户端上传活冷标志位 F_i 至服务端。

需要说明的是，对于冷活标志位 F_i 未发生改变神经元，可以不传该神经元的冷活标志位 F_i 。如果所有神经元的冷活标志位 F_i 均未发生改变，则该 403 的过程可以省略。

405，服务端汇聚各客户端本地局部模型，以及活冷标志位，进行各本地模型的还原。

服务端汇聚各个客户端的本地局部模型 M_I ，其中， I 为客户端数量，如客户端 1 的本地局部模型为 M_{10} ，客户端 2 的本地局部模型为 M_{20} ，客户端 3 的本地局部模型为 M_{30} 等。

例如，由左至右、由上至下对每个神经元排列，并由上至下对每个神经元的相关参数排列，对活冷标志位 $F_i=1$ 的神经元，以最近的中心模型参数补齐，对活冷标志位 $F_i=0$ 的神经元，以本地局部模型参数更新。

406，服务端对各本地模型进行联合平均，更新中心模型。

407，服务端下发中心模型给各客户端，转 401。

重复上述 401 至 407 的过程，直到中心模型收敛，得到收敛的目标 AI 模型，则联合学习训练结束。

针对客户端执行用于获取人工智能模型的方法，以如图 5 所示为例，对客户端进行模型训练的过程进行举例说明。如图 5 所示，包括但不限于如下步骤：

501，客户端接收服务端发送的中心模型。

其中，该中心模型一般为离线训练好的已知模型，神经元数量为 N 。此外，客户端确定本轮训练 E 批次。每轮训练的批次可以由服务端下发至客户端，也可以双方约定。一般情况下，每个客户端每轮训练的批次数量相同。

502，客户端初始化中心模型的神经元的冷冻期、活冷标志位及冷冻期计数器。

例如，初始化中心模型的神经元的冷冻期都为 1，即 $T_i = 1$, $i \in \{1,2,\dots,5\}$ 。若 $T_i=1$ ，则本回本地模型的第 i 个神经元需要进行冷冻探测机制来更新冷冻期。

初始化各神经元的活冷标志位都为 0，即 $F_i = 0$, $i \in \{1,2,\dots,5\}$ 。若 $F_i=1$ ，则本回本地模型的第 i 个神经元被冷冻，即该神经元对应参数不参与训练。

初始化每个神经元的冷冻期计数器 $T_{ci}=1$ 。

503，客户端以初始中心模型为基础，查询各个神经元的 F_i 值，将 $F_i=1$ 的所有神经元冷冻，开始训练 $F_i=0$ 的所有神经元 E 批次。

504，客户端查询各个神经元的 T_i 值，对 $T_{ci}=T_i$ 的所有神经元，启动冷冻探测机制，更新 T_i 、 F_i ，并重新初始化为 $T_{ci}=1$ 。

对 $T_{ci} \neq T_i$ 的所有神经元，保持 T_i 、 F_i ，更新冷冻期内计数器 $T_{ci}=T_{ci}+1$ 。

505，训练结束后，客户端获得本地模型，将本地模型上传给服务端。

本申请实施例提供了一种用于获取人工智能模型的装置，该装置用于执行上述图 3 所示实施例中客户端所执行的方法。如图 6 所示，该装置包括：

通信单元 601，用于接收服务端发送的第一 AI 模型，第一 AI 模型包括多个神经元。例如，该通信单元 601 所执行的功能可如图 3 中的 303 所示，此处不再赘述。

处理单元602, 用于在多个神经元中确定参与本轮训练的目标神经元, 本轮训练为非首轮训练, 目标神经元的数量小于多个神经元的总数量。例如, 该处理单元602所执行的功能可如图3中的303所示, 此处不再赘述。

处理单元602, 还用于基于本地数据对目标神经元进行训练, 得到目标神经元对应的参数。例如, 该处理单元602所执行的功能可如图3中的305所示, 此处不再赘述。

通信单元601, 还用于向服务端返回目标神经元对应的参数数据, 目标神经元对应的参数数据用于服务端获取收敛的目标AI模型。例如, 该通信单元601所执行的功能可如图3中的306所示, 此处不再赘述。

在示例性实施例中, 多个神经元中的各个神经元具有对应的冷活标志位, 冷活标志位的值为第一值时, 用于指示神经元参与本轮训练;

处理单元602, 用于将冷活标志位为第一值的神经元确定为参与本轮训练的目标神经元。

在示例性实施例中, 多个神经元中的各个神经元还具有对应的冷冻期, 冷冻期用于指示神经元不参与训练的周期;

任一神经元对应的冷活标志位的值, 基于任一神经元对应的冷冻期确定。

在示例性实施例中, 处理单元602, 还用于对于多个神经元中的任一神经元, 获取任一神经元在上轮训练的活跃度, 活跃度用于指示神经元被本地数据影响的程度; 根据任一神经元在上轮训练的活跃度更新任一神经元的冷冻期, 根据任一神经元的冷冻期更新任一神经元的冷活标志位的值。

在示例性实施例中, 多个神经元中的各个神经元还具有对应的冷冻期计数器, 冷冻期计数器用于指示冷冻期是否结束;

处理单元602, 用于响应于任一神经元的冷冻期计数器的值指示任一神经元的冷冻期结束, 获取任一神经元在上轮训练的活跃度。

在示例性实施例中, 处理单元602, 用于响应于本轮训练为非首轮训练, 对于多个神经元中的任一神经元, 获取任一神经元在上轮训练的活跃度, 活跃度用于指示神经元被本地数据影响的程度; 响应于任一神经元在上轮训练的活跃度大于活跃度阈值, 将任一神经元确定为参与本轮训练的目标神经元。

在示例性实施例中, 处理单元602, 用于获取任一神经元在上轮训练前得到的参数的第一平均值, 以及任一神经元在上轮训练后得到的参数的第二平均值; 获取第一平均值及第二平均值之间的差值; 根据差值的绝对值与第一平均值的绝对值确定任一神经元在上轮训练的活跃度。

在示例性实施例中, 处理单元602, 还用于接收服务端发送的第一人工智能AI模型, 将多个神经元中的全部神经元确定为目标神经元, 基于本地数据对目标神经元进行训练。

在示例性实施例中, 通信单元601, 还用于向服务器发送各个神经元的冷活标志位, 或者, 向服务器发送值发生变化的冷活标志位。

在示例性实施例中, 通信单元601, 用于仅向服务端返回目标神经元对应的参数数据。

本申请实施例提供的装置, 由于并不需要对模型中的所有神经元均进行训练, 而是按照活跃条件来选择部分神经元进行训练, 从而能够降低功耗, 加快训练速度。此外, 又由于客户端向服务端仅传输训练的神经元对应的参数, 因而能够节省通信带宽。

另外, 通过冷活标志位来指示活跃条件, 从而通过该冷活标志位即可直接确定目标神经

元,使得确定目标神经元的速度较快,进一步节省了训练时间,加快了训练速度。通过设置冷冻期,基于冷冻期来指示神经元不参与训练的周期,基于冷冻周期来确定神经元是否参与训练,从而防止丢失在训练过程中需要参与训练的神经元对应的参数,保证了鲁棒性。通过设置冷冻期计数器,基于冷冻期计数器来指示冷冻期是否结束,并在冷冻期结束后,再获取活跃度,进一步节省了探测冷冻期是否结束的时间,提高了训练速度。

本申请实施例提供了一种用于获取人工智能模型的装置,该装置用于执行上述图3所示实施例中服务端所执行的功能。如图7所示,该装置包括:

处理单元701,用于获取待训练的第一人工智能AI模型,第一AI模型包括多个神经元。例如,该处理单元701所执行的功能可如图3中的301所示,此处不再赘述。

通信单元702,用于向多个客户端发送第一AI模型。例如,该通信单元702所执行的功能可如图3中的302所示,此处不再赘述。

通信单元702,还用于接收多个客户端中的各个客户端返回的目标神经元对应的参数数据,任一客户端返回的目标神经元对应的参数数据由任一客户端对第一AI模型中的目标神经元进行训练得到,目标神经元的数量小于多个神经元的总数量。例如,该通信单元702所执行的功能可如图3中的307所示,此处不再赘述。

处理单元701,还用于根据各个客户端返回的目标神经元对应的参数数据,还原各个客户端对应的第二AI模型,根据各个客户端对应的第二AI模型获取收敛的目标AI模型。例如,该处理单元701所执行的功能可如图3中的307所示,此处不再赘述。

在示例性实施例中,处理单元701,用于对于任一客户端,根据任一客户端返回的目标神经元对应的参数,更新第一AI模型中的目标神经元对应的参数,将第一AI模型中的非目标神经元对应的参数进行补齐,得到任一客户端对应的第二AI模型。

在示例性实施例中,处理单元701,用于对于任一客户端,根据任一客户端返回的目标神经元对应的参数数据,更新第一AI模型中的目标神经元对应的参数,将第一AI模型中的非目标神经元对应的参数进行补齐,得到任一客户端对应的第二AI模型。

在示例性实施例中,处理单元701,用于对各个客户端对应的第二AI模型进行联合平均,得到第三AI模型;响应于第三AI模型收敛,将第三AI模型作为目标AI模型;响应于第三AI模型未收敛,向多个客户端发送第三AI模型,采用获取第三AI模型的方式继续获取新的AI模型,依此循环,直至得到收敛的目标AI模型。

在示例性实施例中,通信单元702,还用于接收多个客户端返回的各个神经元的冷活标志位或者值发生变化的冷活标志位;

处理单元701,用于通过各个神经元的冷活标志位或者值发生变化的冷活标志位确定多个神经元中的目标神经元。

本申请实施例提供的装置,由于客户端并不是对模型中的所有神经元均进行训练,而是按照活跃条件来选择部分神经元进行训练,仅上传训练的目标神经元对应的参数,因而能够节省通信带宽,且提高训练速度,也正因为如此,进而提高了获取目标AI模型的速度。

另外,除了更新第一AI模型中的目标神经元对应的参数,还将第一AI模型中的非目标神经元对应的参数进行补齐,保证了模型的完整性,使得获取的目标AI模型的性能更好。

应理解的是，上述图 6 或图 7 提供的装置在实现其功能时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的装置与方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

图 8 为本申请实施例的用于获取人工智能模型的设备 1100 的硬件结构示意图。图 8 所示的用于获取人工智能模型的设备 1100 可以执行上述实施例的方法中服务端或客户端执行的相应步骤。

如图 8 所示，用于获取人工智能模型的设备 1100 包括处理器 1101、存储器 1102、接口 1103 和总线 1104。其中接口 1103 可以通过无线或有线的方式实现，具体来讲可以是网卡。上述处理器 1101、存储器 1102 和接口 1103 通过总线 1104 连接。

接口 1103 可以包括发送器和接收器，用于获取人工智能模型的设备可通过接口 1103 接收和发送目标神经元对应的参数。

例如，图 8 所示的用于获取人工智能模型的设备 1100 为图 3 中的客户端，处理器 1002 读取存储器 1001 中的指令，使图 8 所示的用于获取人工智能模型的设备 1100 能够执行客户端所执行的全部或部分操作。

又例如，图 8 所示的用于获取人工智能模型的设备 1100 为图 3 中的服务端，处理器 1002 读取存储器 1001 中的指令，使图 8 所示的用于获取人工智能模型的设备 1100 能够执行服务端所执行的全部或部分操作。

另外，处理器 1101 和/或用于本文所描述的技术的其他过程。存储器 1102 包括操作系统 11021 和应用程序 11022，用于存储程序、代码或指令，当处理器或硬件设备执行这些程序、代码或指令时可以完成方法实施例中涉及用于获取人工智能模型的设备 1100 的处理过程。可选的，存储器 1102 可以包括只读存储器（英文：Read-only Memory，缩写：ROM）和随机存取存储器（英文：Random Access Memory，缩写：RAM）。其中，ROM 包括基本输入/输出系统（英文：Basic Input/Output System，缩写：BIOS）或嵌入式系统；RAM 包括应用程序和操作系统。当需要运行用于获取人工智能模型的设备 1100 时，通过固化在 ROM 中的 BIOS 或者嵌入式系统中的 boot loader 引导系统进行启动，引导用于获取人工智能模型 1100 进入正常运行状态。在用于获取人工智能模型的设备 1100 进入正常运行状态后，运行在 RAM 中的应用程序和操作系统，从而，完成方法实施例中涉及用于获取人工智能模型的设备 1100 的处理过程。

可以理解的是，图 8 仅仅示出了用于获取人工智能模型的设备 1100 的简化设计。在实际应用中，用于获取人工智能模型的设备 1100 可以包含任意数量的接口，处理器或者存储器。

应理解的是，上述处理器可以是中央处理器（Central Processing Unit，CPU），还可以是其他通用处理器、数字信号处理器（digital signal processing，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、分立硬件组件等。通用处理器可以是微处理器或者是任何常规的处理器等。值得说明的是，处理器可以是支持进阶精简指令集机器（advanced RISC machines，ARM）架构的处理器。

进一步地，在一种可选的实施例中，上述存储器可以包括只读存储器和随机存取存储器，

并向处理器提供指令和数据。存储器还可以包括非易失性随机存取存储器。例如，存储器还可以存储设备类型的信息。

该存储器可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器（read-only memory, ROM）、可编程只读存储器（programmable ROM, PROM）、可擦除可编程只读存储器（erasable PROM, EPROM）、电可擦除可编程只读存储器（electrically EPROM, EEPROM）或闪存。易失性存储器可以是随机存取存储器（random access memory, RAM），其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的RAM可用。例如，静态随机存取存储器（static RAM, SRAM）、动态随机存取存储器（dynamic random access memory, DRAM）、同步动态随机存取存储器（synchronous DRAM, SDRAM）、双倍数据速率同步动态随机存取存储器（double data rate SDRAM, DDR SDRAM）、增强型同步动态随机存取存储器（enhanced SDRAM, ESDRAM）、同步连接动态随机存取存储器（synchlink DRAM, SLDRAM）和直接内存总线随机存取存储器（direct rambus RAM, DR RAM）。

还提供了一种计算机可读存储介质，存储介质中存储有至少一条指令，指令由处理器加载并执行以实现如上任一所述的获取人工智能模型的方法。

本申请提供了一种计算机程序，当计算机程序被计算机执行时，可以使得处理器或计算机执行上述方法实施例中对应的各个步骤和/或流程。

提供了一种芯片，包括处理器，用于从存储器中调用并运行所述存储器中存储的指令，使得安装有该芯片的通信设备执行上述各方面中的方法。

提供另一种芯片，包括：输入接口、输出接口、处理器和存储器，所述输入接口、输出接口、所述处理器以及所述存储器之间通过内部连接通路相连，所述处理器用于执行所述存储器中的代码，当所述代码被执行时，所述处理器用于执行上述各方面中的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线）或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘 Solid State Disk）等。

以上所述的具体实施方式，对本申请的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本申请的具体实施方式而已，并不用于限定本申请的保护范围，凡在本申请的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本申请的保护范围之内。