

预训练服务系统及基于预训练服务系统的服务提供方法

5 本申请要求于 2022 年 04 月 06 日提交中国专利局、申请号为 202210353465.X、
申请名称为“预训练服务系统及基于预训练服务系统的服务提供方法”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

10 本申请一个或多个实施例涉及深度学习领域，尤其涉及一种预训练服务系统及基
于预训练服务系统的服务提供方法。

背景技术

15 在人工智能技术飞速发展的背景下，各种大规模预训练模型应运而生，预训练模
型所具有的理解、生成和检索等能力已被应用至各行各业并取得了显著成功。但是，
由于预训练模型的使用门槛较高，如何降低使用难度，为更多用户提供通过简单操作
就可以运用预训练模型的平台，是亟待解决的问题。并且，由于预训练模型具有极强
的理解、生成和检索能力，如何避免预训练模型的滥用也是考虑的问题。

发明内容

20 有鉴于此，本申请一个或多个实施例提供一种预训练服务系统。
为实现上述目的，本申请一个或多个实施例提供技术方案如下：
根据本申请一个或多个实施例的第一方面，提出了一种预训练服务系统，所述预
训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，所述已混淆
模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进
行相关联的混淆处理后生成；所述预训练服务系统包括：
25 生产者服务模块，用于向模型生产者提供针对预训练数据集的模型预训练功能，
并产出相应的预训练模型；
优化者服务模块，用于根据模型优化者提供的微调数据集对所述预训练模型进行
优化，并得到优化后模型；
消费者服务模块，用于向模型消费者提供针对所述预训练模型和/或所述优化后模
30 型的接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数
据进行推理并输出模型预测结果。
根据本申请一个或多个实施例的第二方面，提出了一种基于预训练服务系统的服
务提供方法，所述预训练服务系统包括生产者服务模块、优化者服务模块和消费者服
务模块；所述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参
35 数，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和
原始模型参数进行相关联的混淆处理后生成；所述方法包括：
获取访问所述预训练服务系统的目标用户的身份信息，以确定所述目标用户所属

的预设用户类型；

向所述目标用户开放与其所属的预设用户类型相匹配的服务模块，其中：

在所述目标用户为模型生产者的情况下，向所述目标用户开放所述生产者服务模块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；

在所述目标用户为模型优化者的情况下，向所述目标用户开放所述优化者服务模块，所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型；

在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

根据本申请一个或多个实施例的第三方面，提出了一种基于预训练服务系统的服务提供装置，所述预训练服务系统包括生产者服务模块、优化者服务模块和消费者服务模块；所述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成；所述装置包括：

获取单元，用于获取访问所述预训练服务系统的目标用户的身份信息，以确定所述目标用户所属的预设用户类型；

开放单元，用于向所述目标用户开放与其所属的预设用户类型相匹配的服务模块，其中：

在所述目标用户为模型生产者的情况下，向所述目标用户开放所述生产者服务模块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；

在所述目标用户为模型优化者的情况下，向所述目标用户开放所述优化者服务模块，所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型；

在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

根据本申请一个或多个实施例的第四方面，提出了一种电子设备，包括：处理器、内存及非易失性存储器；

所述处理器用于从所述非易失性存储器中读取对应的计算机程序到所述内存中，并基于所述计算机程序执行如上第二方面所述的方法。

根据本申请一个或多个实施例的第五方面，提出了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如上第二方面所述的方法。

根据本申请一个或多个实施例的第六方面，提出了一种计算机程序产品，包括计

算机程序，当所述计算机程序被运行时，使得计算机执行如上第二方面所述的方法。

在本申请所提供的技术方案中，通过一种包括生产者服务模块、优化者服务模块和消费者服务模块的预训练服务系统为具有不同身份的用户提供不同类型的服务，降低了预训练模型的使用难度，提高了预训练模型的普适性，为不具有生成预训练模型的能力的模型优化者提供快速便捷地基于预训练模型获取优化后模型的能力，同时，还提供服务接口给模型消费者，使模型消费者可以通过服务接口快速利用获取预训练模型或者优化后模型进行推理并获取模型预测结果。且由于预训练服务系统提供的模型结构代码以及模型参数为已混淆模型结构代码和已混淆模型参数，其在功能上与原始模型结构代码及原始模型参数等价，但是通过已混淆模型结构代码难以复现该预训练模型，使用已混淆结构代码和已混淆模型参数可以避免预训练模型的滥用，从而达到保护预训练模型的效果。

附图说明

图 1 是本申请一示例性实施例提供的一种预训练服务系统的结构示意图；
图 2 是本申请一示例性实施例提供的一种基于预训练服务系统的服务提供方法的流程图示意图；
图 3 是本申请一示例性实施例提供的一种电子设备的结构示意图；
图 4 是本申请一示例性实施例提供的一种基于预训练服务系统的服务提供装置示意图。

具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请一个或多个实施例相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本申请一个或多个实施例的一些方面相一致的装置和方法的例子。

需要说明的是：在其他实施例中并不一定按照本申请示出和描述的顺序来执行相应方法的步骤。在一些其他实施例中，其方法所包括的步骤可以比本申请所描述的更多或更少。此外，本申请中所描述的单个步骤，在其他实施例中可能被分解为多个步骤进行描述；而本申请中所描述的多个步骤，在其他实施例中也可能被合并为单个步骤进行描述。

人工智能（Artificial Intelligence，AI）技术的应用场景的复杂化导致在不同的应用场景中，任务目标、训练数据、问题设定等方面均存在差异。因此，通过向用户提供具有通用性的预训练模型，以使用户在该预训练模型的基础上进行微调以获得符合特定应用场景的优化后模型的应用模式降低了对训练数据的需求，提升了训练效率，使得更多的中小型企业具备了获得符合特定应用场景的优化后模型的能力，从而推进了 AI 技术的应用普及。同时，用户还可以根据应用场景选取不同的预训练模型或者服务特定场景的优化后模型利用自己的数据进行推理，从而获取基于选定的预训练模型或者优化后模型的模型输出结果。

为了方便为具有不同身份类型的用户提供不同种类的基于预训练模型的服务，本申请一示例性实施例提供一种如图 1 所示的预训练服务系统，该预训练服务系统包括：生产者服务模块 11、优化者服务模块 12 和消费者服务模块 13。其中，生产者服务模块 11 用于向模型生产者提供针对预训练数据集 111 的模型预训练功能，并产出相应的预训练模型 112；优化者服务模块 12，用于根据模型优化者提供的微调数据集 121 对所述预训练模型 112 进行优化，并得到优化后模型 122；消费者服务模块 13，用于向模型消费者提供针对所述预训练模型 112 和/或所述优化后模型 122 的服务接口，所述预训练模型 112 或所述优化后模型 122 用于对所述模型消费者提供的的数据 131 进行推理并输出模型预测结果 132。

其中，上述模型生产者是指预训练服务系统方的算法开发人员，是预训练模型 112 的生产者，在本申请一示例性实施例中，模型生产者基于预训练数据集 111 利用预训练服务系统提供的模型预训练功能启动分布式预训练并产出预训练模型 112，该预训练模型 112 是预训练服务系统提供人工智能相关服务的基础，其他预训练模型均为由该预训练模型 112 经修改后获得。

在本申请一示例性实施例中，在预训练模型 112 的预训练过程中，模型预训练功能实现了激活检查点（Activation Checkpointing）技术。在网络中设置若干个检查点（Checkpoint），并在前向过程中舍弃部分暂时无用的中间结果以降低内存占用，若在后续过程中需要这些中间结果时，则从最近的检查点通过额外的前向计算恢复对应的中间结果，这样既节省了显存，又避免了从头计算的繁琐过程，从而降低预训练模型训练过程中的中间激活（Activation）带来的显存占用。

此外，预训练过程中还采用了混合精度训练技术，通过 16 位浮点数（FP16）进行深度学习预训练模型的训练以降低训练所需的显存，提高计算的速度。同时，预训练系统还采用稀疏专家并行和数据并行等技术提高预训练过程的计算效率，保证了预训练过程的低碳环保。其中，稀疏专家并行技术是指稀疏门控专家混合网络（Mixture of Experts, MoE），将大模型拆分成多个小模型，每个小模型保留一份独立的参数，模型预训练过程中，对于一个样本来说，无须经过所有的小模型去计算，而是通过预设的策略选择适当的小模型，选取对应的参数进行计算，这种方案中预训练总参数量虽然很高，但是被激活的只有一小部分，节省了计算过程中的显存使用。

在本申请一示例性实施例中，为了使得产出的预训练模型 112 所具有的能力更加多元化，且具有更高的通用性，在预训练模型 112 的预训练过程中，采取多模态预训练的形式。其中，模态是指数据的存在形式，比如文本、音频、图像、视频等文件格式。

在本申请一示例性实施例中，在预训练模型 112 的预训练过程中，预训练模型 112 可以使用包括：语言模型（Language Modeling），文本去噪（Text Denoising），图像描述生成（Image Captioning），基于图像的文本去噪（Image-based Text Denoising）等多个任务。此外，在预训练过程中，预训练数据集 111 的规模和领域广泛程度影响了最后获得的预训练模型 112 的能力以及通用性。因此，在本申请一示例性实施例中，预训练模型 112 的预训练过程使用了覆盖多个领域的超大规模单模态及多模态数据集作为预训练模型 112 的预训练数据集 111。通过启动多模态的预训练以及扩大预训练

数据集 111 的规模及覆盖领域可以使得预训练模型 112 同时具备单模态及多模态的理解与生成等能力,而且在不同领域任务上都有良好的适应能力,提高了预训练模型 112 的通用性。

对应于模型开发者,模型优化者是指其他领域中,想基于预训练模型 112 获得适配于自己所属的领域的优化后模型 122 的用户。在本申请一示例性实施例中,如图 1 所示的预训练服务系统为服务优化者提供优化者服务模块 12,上述优化者服务模块 12 根据模型优化者提供的微调数据集 121 对由生产者服务模块 11 产出的预训练模型 112 进行优化,并得到优化后模型 122。

在本申请一示例性实施例中预训练服务系统提供的模型预训练功能用于根据预训练数据集 111 训练得到通用预训练模型。然后生产者服务模块根据上述通用预训练模型生成对应于模型生产者所指示的下游任务的专用预训练模型。即模型生产者根据通用预训练模型原子能力指示预训练服务系统,在通用预训练模型的基础上生成具有专门用途的、适用于某种类型的下游任务的专用预训练模型。其中通用预训练模型的原子能力包括跨模态生成能力、跨模态理解能力和跨模态检索能力等。相对应的,专用预训练模型可以包括:跨模态生成模型、跨模态理解模型和跨模态检索模型等。预训练模型 112 包括:通用预训练模型和专用预训练模型。

在本申请一示例性实施例中,预训练服务系统中的生产者服务模块 11 还用于生成专用预训练模型对应的代码开发模板。上述专用预训练模型对应的代码开发模板由预训练服务系统通过优化者服务模块 12 提供给模型优化者。

任一预训练模型主要包括两部分:模型结构代码和模型参数。其中,模型结构代码是指构建预训练模型的代码,可以将模型看作假设,而模型参数是根据特定的数据集对假设进行的具体调整,而模型参数值可以定义模型功能。

模型优化者参考上述代码开发模板,根据自身的优化需求,向预训练服务系统发出模型修改指令。具体地,预训练服务系统中的优化者服务模块 12 为模型优化者提供模型修改接口,上述模型优化者参考代码开发模板利用所述模型修改接口针对专用预训练模型的模型结构代码和/或模型参数发出模型修改指令,优化者服务模块根据模型修改指令修改对应的模型结构代码和/或模型参数,然后针对修改后的预训练模型根据模型优化者提供的微调数据集 121 启动训练,获取优化后模型。

在本申请一示例性实施例中,对于想快速试用预训练模型的模型优化者,只需通过修改上述代码开发模板提供示例的参数,即可利用本领域的数据(微调数据集),通过预训练服务系统启动多机多卡的分布式训练任务并产出属于该模型优化者的优化后模型参数。将该优化后模型参数与预训练模型的模型结构代码进行结合即可产出优化后模型。

在本申请另一示例性实施例中,若模型优化者想进一步对预训练模型进行优化,则预训练服务系统还可以提供将预训练模型优化并运用于其他领域的示范例,优化者可以根据自己的需求,参考上述示范例对预训练模型下发修改指令,并输入本领域的微调数据集,则预训练服务系统启动分布式训练任务,产出对应的优化后模型参数,从而获取优化后模型。

该优化后模型基于模型优化者提供的微调数据集进行训练,且模型优化者可以根

据自身需求对模型进行优化微调，因此，优化后模型适用于模型优化者所属领域。且在利用预训练服务系统进行模型优化的过程中，模型优化者无需关注预训练模型的开发过程、预训练模型的分布式训练、预训练模型的评估流程等的具体实现细节，例如，数据 IO、数据并行等，仅需要根据预训练服务系统提供的代码开发模板修改对应的模型结构代码或者模型参数并启动模型训练，即可获取适用于模型优化者所属行业的优化后模型。预训练服务系统降低了用户优化预训练模型的难度，使得不具有预训练模型开发能力的模型优化者也可以通过预训练服务系统优化后模型。

在本申请一示例性实施例中，对于已经趋向成熟的行业解决方案（如文案生成、搜索召回、服饰设计、智能对话、金融问答等），可以将其作为示范例由预训练服务系统提供给其他用户（包括其他开发者）进行参考。模型优化者除参考代码开发模板对预训练模型发送修改指令外，还可以参考行业解决方案对模型进行修改。

在本申请一示例性实施例中，预训练服务系统还提供分享接口，模型优化者可以将对预训练模型的优化过程通过分享接口上传至预训练服务系统，作为新的示范例进行分享，提供给其他模型优化者作为优化参考，通过案例分享功能可以扩大预训练模型的应用范围，进一步降低预训练服务系统的使用难度。

在本申请中，预训练服务系统的用户还包括模型消费者，模型消费者是想通过预训练模型或者优化后模型利用属于自己的数据进行推理，以获取模型预测结果的用户。预训练服务系统通过消费者服务模块向模型消费者提供针对预训练模型和/或优化后模型的服务接口，预训练模型或优化后模型用于对模型消费者提供的数据进行推理并输出模型预测结果。如图 1 所示，模型消费者可以调用预训练模型 112 或者优化后模型 122 进行推理。上述推理过程可以是离线的，预训练服务系统通过服务接口将模型消费者选中的预训练模型 112 或优化后模型 122 部署于该模型消费者对应的设备，以用于对该模型消费者提供的数据进行离线推理。当然，上述预训练模型 112 或者优化后模型 122 也可以部署为低延时的在线服务，消费者通过在预训练服务系统提供的 web 页面内直接输入属于自己的数据 131，该 web 页面相当于预训练服务系统提供的服务接口，获取模型消费者的数据 131 后，调用对应的预训练模型 112 或优化后模型 122 进行在线推理，并得到模型预测结果，将该模型预测结果通过 web 页面的形式提供给模型消费者。

在本申请一示例性实施例中，预训练服务系统可以基于预训练模型的跨模态生成、跨模态检索、跨模态理解能力为用户提供单模态或跨模态的多种不同的人工智能服务。

例如，对于某一新闻事件，可以通过文本、图像或者视频等多种形式展示，因此，在进行信息检索时，如果仅针对单一模态的数据进行检索，获取的信息内容将受到很大的限制，需要其他模态的数据来丰富我们对同一事物或事件的认知，此时就需要跨模态检索来实现不同模态数据之间的检索。在本申请一示例性实施例中，预训练服务系统可基于预训练模型的跨模态检索能力，根据用户输入的图片形式的新闻信息获取该新闻事件的视频。

在本申请一示例性实施例中，预训练服务系统可基于预训练模型的跨模态生成能力为模型消费者提供服装设计的服务，例如，模型消费者通过预训练服务系统提供的 web 页面，输入“男士条纹衬衫”，通过对预训练服务系统提供的服装设计领域的目

标预训练模型的计算，可以生成若干符合“男士条纹衬衫”文字描述的服饰设计图片。上述示例利用了预训练模型的跨模态生成能力。当然，虽然预训练服务系统为用户提供跨模态的人工智能服务，但是预训练服务系统也具有基于单模态提供人工智能服务的能力，例如，当消费者通过预训练服务系统提供的服饰设计领域的目标预训练模型的离线推理功能进行如下的服饰设计操作时，模型消费者输入两张不同样式的衣物图片，进行离线推理后，模型消费者可以获得离线推理的结果——将两张样式不同的衣物进行融合后获得的衣物图片。

除了服饰设计功能之外，预训练服务系统还可以利用预训练模型的跨模态检索、生成、理解能力进行诸如文案生成等与文字处理相关的训练推理。例如，在对文字处理领域的目标预训练模型的推理过程中，输入一张乳胶枕头的图片，利用预训练服务系统提供的文字处理领域的目标预训练模型进行推理时，利用该预训练模型的跨模态理解能力，可以获得对于该乳胶枕头的宣传文案。另外，该预训练服务系统还利用预训练模型的上述能力实现了从文本到文本的转换，并在转换过程中，执行了文本去噪和语言建模。

通过上述对预训练服务系统的介绍可知，在预训练服务系统的服务过程中，预训练模型具有十分重要的作用，用户可以通过预训练服务系统调用目标预训练模型，而由于预训练模型具有极强的文本和图像生成能力，为避免预训练模型的滥用，预训练服务系统还设置有多种防止预训练模型滥用的防御机制，且该防御机制同样适用于避免优化后模型的滥用。

通过上文介绍可知，预训练模型或者优化后模型包括两部分：模型结构代码和模型参数。对于模型结构代码，根据行业通用的框架等资料，用户有可能逆向复现出完整的模型结构代码，并通过将复现的模型结构代码以及通过其他手段获取的模型参数相结合从而获取到预训练模型或者优化后模型。不利于对预训练模型或者优化后模型应用的管控。而由于模型参数定义了模型的功能，因此，对于模型参数的保护和模型结构代码的保护对于预训练服务系统同样重要。在对预训练模型整体进行保护时，不仅需要考虑对模型结构代码的保护，也要考虑对模型参数的保护。

在本申请一示例性实施例中，预训练模型和/或优化后模型包括：已混淆模型结构代码和已混淆模型参数；其中，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成。通过将模型结构代码和模型参数进行关联混淆，使得已混淆模型结构代码及已混淆模型参数虽然具有与原始模型结构代码与原始模型参数等价的功能，但是在形式上无法被复现，从而达到避免预训练模型被滥用的效果。模型参数值可以定义模型功能。

具体地，该相关联的混淆处理可以以如下形式实现：在对预训练模型的原始模型结构代码中的至少一部分代码进行常规混淆操作（通常通过移除注释，注入无关代码，重命名函数、类、变量等方法降低代码可读性，提高破解门槛，但保证运行结果上与混淆前一致）的基础上，对模型结构代码中至少一处计算过程进行修改，同时配合修改模型参数中与该计算过程相应的部分参数，以使经修改的已混淆模型结构代码与进行过相应修改的已混淆模型参数结合而构成的已混淆预训练模型的运行结果与混淆前的原始预训练模型相一致。由于该混淆过程配合了对模型参数的修改调整，因此，他

人难以定位发生混淆的具体位置，上述方法进一步提高了复现模型结构代码的难度，实现了对模型结构代码的保护。

前述对模型结构代码中至少一处计算过程进行修改，包括：对未经修改的原始模型结构代码中的至少一部分代码进行正向混淆处理，配合对该原始模型参数中与进行修改的部分代码相关的模型参数进行逆向混淆处理，在不改变预训练模型运行结果的基础上，增大预训练模型的复现难度。

例如，前述正向混淆处理可以包括：在对预训练模型的原始模型结构代码中的至少一部分代码进行常规混淆操作的基础上，在模型结构代码的计算过程中减去某个向量。而对应的逆向混淆处理可以包括：在模型参数中对应部分加上这个向量进行补偿，以使已混淆预训练模型的运行结果与混淆前的原始预训练模型相一致。

为进一步提高预训练模型的安全性，在本申请一示例性实施例中，将模型参数存储在预训练服务系统管理的存储空间中，当目标用户针对目标预训练模型或目标优化后模型进行调用时，根据目标用户的调用请求中指示的读取路径，从该存储空间中获取相应的模型参数。其中，该存储空间可以为受预训练服务系统管理的对象存储（云存储），或者，该存储空间也可以直接由预训练服务系统提供。上述模型参数可以是已混淆模型参数。在预训练服务系统获取该已混淆模型参数后，将其与对应的已混淆模型结构代码进行结合以获取目标预训练模型或者目标优化后模型。

由于预训练服务系统的用户具有不同的身份，包括：模型开发者、模型优化者和模型消费者。不同类型的用户具有不同的需求，其利用预训练服务系统完成不同的任务，因此，若在预训练服务系统上向全部类型的用户开放全部的使用权限，将不利于对预训练模型的保护。

为根据用户身份提供相应的模型使用权限，本申请中的预训练服务系统还提供如下的服务：

优化者服务模块还用于：鉴定所述模型优化者是否具有调用所述预训练模型对应的已混淆模型参数的权限；若所述模型优化者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型并提供至所述模型优化者；和/或，

所述消费者服务模块还用于：鉴定所述模型消费者是否具有调用所述预训练模型或所述优化后模型对应的已混淆模型参数的权限；若所述模型消费者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型或所述优化后模型并提供至所述模型消费者。

例如，预训练服务系统中的优化者服务模块鉴定某模型优化者具有调用其选择的目标预训练模型的模型参数的权限，其中该模型参数可以为已混淆模型参数。若该模型优化者具有调用上述已混淆模型参数的权限，则在系统获取该已混淆参数模型后，与对应的已混淆模型结构代码进行结合，即可获得目标预训练模型。将该目标预训练模型提供给上述模型优化者以供其对该目标预训练模型进行优化。对于模型消费者，鉴定过程与上述模型优化者类似。

当然，在预训练服务系统的运行过程中，除了预训练模型和优化后模型的模型结构代码和模型参数外，为了对预训练模型或者优化后模型进行调取和使用，预训练服

务系统中还运行有针对任一预训练模型或优化后模型的模型存取代码和模型训练推理代码，将此部分代码与模型结构代码进行解耦，获得独立的模型存取代码和模型训练推理代码，并将模型存取代码与模型结构代码进行加密处理，可以避免预训练模型或者优化后模型的存储位置暴露或者模型结构代码的外泄，提高预训练服务系统的安全性，避免预训练模型或者优化后模型被滥用。但对于模型训练推理代码，则可以由预训练服务系统明文维护。

在本申请一示例性实施例中，提供一种如图 2 所示的基于预训练服务系统的服务提供方法。该方法可以应用于如图 1 所示的预训练服务系统中。上述预训练服务系统可以包括生产者服务模块 11、优化者服务模块 12 和消费者服务模块 13。上述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，已混淆模型结构代码与已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成；上述方法可以包括如下步骤：

S201，获取访问所述预训练服务系统的目标用户的身份信息，以确定所述目标用户所属的预设用户类型。

例如，预训练服务系统可以根据目标用户的登录账号确定该目标用户的身份信息，以确定目标用户所属的预设用户类型。或者，预训练服务系统可以根据目标用户的调用请求确定该目标用户的身份信息。

S202，向所述目标用户开放与其所属的预设用户类型相匹配的服务模块。

其中：在所述目标用户为模型生产者的情况下，向所述目标用户开放所述生产者服务模块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型。

在所述目标用户为模型优化者的情况下，向所述目标用户开放所述优化者服务模块，所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型。

在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

其中，上述模型预训练功能用于根据预训练数据集训练得到通用预训练模型；其中，所述预训练模型包括以下至少之一：通用预训练模型、生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。

在本申请一示例性实施例中，上述基于预训练服务系统的服务提供方法还包括：生成所述专用预训练模型对应的代码开发模板；将所述代码开发模板提供至所述模型优化者；以及，根据所述模型优化者发出的模型修改指令，对所述专用预训练模型的模型结构代码和/或模型参数进行修改。

例如，在本申请一示例性实施例中，假设服装设计领域的模型优化者（目标用户 A）需要基于目标预训练模型进行修改，以获取针对目标领域即服装设计领域的优化后模型，假设该目标预训练模型为如图 1 所示的预训练模型 112。此时，该服装设计领域的目标用户 A 登录预训练服务系统后，目标预训练系统获取该目标用户的身份信

息，并确认该目标用户为模型优化者。预训练服务系统向目标用户 A 开放优化服务模块 12。目标用户 A 可以参考预训练服务系统提供的代码开发模板，对预训练模型 112 的参数进行修改，并利用属于服装设计领域的目标用户 A 自身的微调数据集 121 中的数据对修改后的预训练模型进行训练，以获取优化后模型 122。或者，该目标用户 A 也可以选择利用预训练服务系统赋予的优化权限，通过预训练服务系统提供的修改接口下发修改指令指示预训练服务系统修改对应的模型结构代码，同样，利用微调数据集启动模型训练，并最终获取优化后模型 122。该优化后模型 122 是适用于服装设计领域的。

在本申请另一示例性实施例中，若预训练服务系统确定目标用户 B 的用户类型为模型消费者，则预训练服务系统向目标用户 B 开放消费者服务模块 13，所述消费者服务模块 13 用于向所述模型消费者提供针对预训练模型和/或所述优化后模型的服务接口，其中，预训练模型或优化后模型用于对模型消费者提供的数据进行推理并输出模型预测结果。例如，当目标用户 B 想要通过预训练服务系统提供的人工智能服务获取特定主题的服装设计方案的时，目标用户 B 请求调用服装设计领域内的目标优化后模型——在上一实施例中用开发者目标用户 A 在预训练模型 112 的基础上优化而获得的优化后模型 122。预训练服务系统的消费者服务模块 13 通过服务接口向目标用户 B 提供利用优化后模型 122 进行推理以获取模型预测结果的模型使用权限。具体地，目标用户 B 可以利用预训练服务系统中在线部署的推理服务，输入“男士条纹衬衫”，调用优化后模型 122 进行在线的推理计算，获取模型预测结果，例如，若干符合“男士条纹衬衫”文字描述的服饰设计图片。当然，该推理功能也可以被设置为离线功能，例如，通过预训练服务系统的消费者服务模块提供的服务接口将优化后模型 122 部署于模型消费者——用户 B 对应的设备，以用于对模型消费者 B 提供的数据进行离线推理。

类似地，模型消费者也可以调用如文字处理领域、金融领域等模型优化者以如上述实施例中相同的优化方式获取的不同领域内的优化后模型，进行推理以获取不同领域的模型预测结果。

由于模型优化者通过预训练服务系统对目标预训练模型进行优化后，需要将优化后模型提供给其他用户进行使用，因此，在本申请一示例性实施例中，预训练服务系统还提供如下的功能：将预训练服务系统产出的优化后模型的参数存入由该预训练服务系统管理的存储空间。上传后的优化后模型参数可以通过上文所述的方式，根据目标用户提供的读取路径被调用。例如，如上文所述实施例中，目标用户 A 获得优化后模型 122 后将其模型参数存入预训练服务系统管理的存储空间。当目标用户 B 想要调用该优化后模型 122 时，预训练服务系统可以通过用户 B 提供的读取路径，调用该优化后模型 122 的模型参数。当然，上述模型参数可以为已混淆模型参数，将该已混淆模型参数与对应的已混淆模型结构代码相结合即可获得优化后模型 122。其中，上述已混淆模型结构代码与已混淆模型参数为预先针对优化后模型 122 的原始模型结构代码和原始模型参数进行相关联的混淆处理后生成。上述相关联的混淆处理的具体方法请参考上文中关于预训练服务系统的相关介绍。

在本申请一示例性实施例中，预训练服务系统可以鉴定目标用户是否具有调用已

混淆模型参数的权限；若所述目标用户具有调用该已混淆模型参数的权限，则将该已混淆模型参数与相应的已混淆模型结构代码相结合以获取目标预训练模型或者目标优化后模型。模型参数通常以 ckpt 文件的形式存储，为了保证模型参数的 ckpt 文件无法脱离预训练服务系统独立使用，可以引入授权服务器（License Server），对目标用户是否具有调用预训练服务系统提供的模型参数（ckpt 文件）的权限进行鉴别，若该目标用户具有上述权限，则将获取到的模型参数与模型结构代码相结合以获取目标预训练模型。

图 3 是本申请一示例性实施例的一种电子设备的结构示意图。请参考图 3，在硬件层面，该设备包括处理器 302、内部总线 304、网络接口 306、内存 308 及非易失性存储器 310。当然还可能包括其他功能所需要的硬件。处理器 302 从非易失性存储器 310 中读取对应的计算机程序到内存 308 中然后运行，在逻辑层面上形成一种基于预训练服务系统的服务提供装置。当然，除了软件实现方式之外，本申请一个或多个实施例并不排除其他实现方式，比如逻辑器件抑或软硬件结合的方式等等，也就是说以下处理流程的执行主体并不限定于各个逻辑单元，也可以是硬件或逻辑器件。

相应于上述方法的实施例，如图 4 所示，本申请还提供一种基于预训练服务系统的服务提供装置，所述预训练服务系统包括生产者服务模块、优化者服务模块和消费者服务模块；所述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成；所述装置包括：

获取单元 410，用于获取访问所述预训练服务系统的目标用户的身份信息，以确定所述目标用户所属的预设用户类型；

开放单元 420，用于向所述目标用户开放与其所属的预设用户类型相匹配的服务模块，其中：

在所述目标用户为模型生产者的情况下，向所述目标用户开放所述生产者服务模块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；

在所述目标用户为模型优化者的情况下，向所述目标用户开放所述优化者服务模块，所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型；

在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

可选地，所述模型预训练功能用于根据所述预训练数据集训练得到通用预训练模型；其中，所述预训练模型包括以下至少之一：所述通用预训练模型、所述生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。

可选地，所述基于预训练服务系统的服务提供装置还可以包括：

生成单元 430，用于生成所述专用预训练模型对应的代码开发模板；

提供及修改单元 440，用于将所述代码开发模板提供至所述模型优化者；以及，根据所述模型优化者发出的模型修改指令，对所述专用预训练模型的模型结构代码和/或模型参数进行修改。

可选地，所述基于预训练服务系统的服务提供装置还可以包括：

- 5 离线部署单元 450，用于通过所述服务接口将所述预训练模型或所述优化后模型部署于所述模型消费者对应的设备，以用于对所述模型消费者提供的数据进行离线推理；或者，

在线推理单元 460，用于通过所述服务接口获取所述模型消费者提供的数据，并调用所述预训练模型或所述优化后模型进行在线推理。

- 10 可选地，所述相关联的混淆处理包括：对所述原始模型结构代码中的至少一部分代码进行正向混淆处理，并对所述原始模型参数中与所述至少一部分代码相关的模型参数进行逆向混淆处理。

可选地，所述已混淆模型参数存储在所述预训练服务系统管理的存储空间中。

可选地，所述基于预训练服务系统的服务提供装置还可以包括：

- 15 第一鉴定单元 470，用于鉴定所述模型优化者是否具有调用所述预训练模型对应的已混淆模型参数的权限；若所述模型优化者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型并提供至所述模型优化者；和/或，

- 20 第二鉴定单元 480，用于鉴定所述模型消费者是否具有调用所述预训练模型或所述优化后模型对应的已混淆模型参数的权限；若所述模型消费者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型或所述优化后模型并提供至所述模型消费者。

- 25 可选地，所述预训练服务系统针对所述目标预训练模型维护有独立的模型存取代码和模型训练推理代码；其中，所述模型结构代码以及所述模型存取代码由所述预训练服务系统经加密后维护，所述模型训练推理代码由所述预训练服务系统明文维护。

上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程，在此不再赘述。

本申请还提供一种计算机程序产品，包括计算机程序，当所述计算机程序被运行时，使得计算机执行如上实施例所述的方法。

- 30 本申请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存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如上实施例所述的方法。

- 上述实施例阐明的系统、装置、模块或单元，具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。一种典型的实现设备为计算机，计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任意几种设备的组合。

在一个典型的配置中，计算机包括一个或多个处理器（CPU）、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器（RAM）和

/或非易失性内存等形式，如只读存储器（ROM）或闪存（flash RAM）。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存储、磁盒式磁带、磁盘存储、量子存储器、基于石墨烯的存储介质或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体（transitory media），如调制的数据信号和载波。

在本申请一个或多个实施例中，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。

在本申请一个或多个实施例使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请一个或多个实施例。在本申请一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

应当理解，尽管在本申请一个或多个实施例可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请一个或多个实施例范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

以上所述仅为本申请一个或多个实施例的较佳实施例而已，并不用以限制本申请一个或多个实施例，凡在本申请一个或多个实施例的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请一个或多个实施例保护的范围之内。

权 利 要 求 书

1. 一种预训练服务系统，其特征在于，所述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成；所述
- 5 预训练服务系统包括：
- 生产者服务模块，用于向模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；
- 优化者服务模块，用于根据模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型；
- 10 消费者服务模块，用于向模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。
2. 如权利要求 1 所述的系统，其特征在于，所述模型预训练功能用于根据所述预训练数据集训练得到通用预训练模型；其中，所述预训练模型包括以下至少之一：所述通用预训练模型、所述生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。
- 15 3. 如权利要求 2 所述的系统，其特征在于，所述生产者服务模块还用于：生成所述专用预训练模型对应的代码开发模板；
- 所述优化者服务模块还用于：将所述代码开发模板提供至所述模型优化者；以及，
- 20 根据所述模型优化者发出的模型修改指令，对所述专用预训练模型的模型结构代码和/或模型参数进行修改。
4. 如权利要求 1-3 任一项所述的系统，其特征在于，所述消费者服务模块用于：
- 通过所述服务接口将所述预训练模型或所述优化后模型部署于所述模型消费者对应的设备，以用于对所述模型消费者提供的数据进行离线推理；或者，
- 25 通过所述服务接口获取所述模型消费者提供的数据，并调用所述预训练模型或所述优化后模型进行在线推理。
5. 如权利要求 1-4 任一项所述的系统，其特征在于，所述相关联的混淆处理包括：对所述原始模型结构代码中的至少一部分代码进行正向混淆处理，并对所述原始模型参数中与所述至少一部分代码相关的模型参数进行逆向混淆处理。
- 30 6. 如权利要求 1-5 任一项所述的系统，其特征在于，所述已混淆模型参数存储在所述预训练服务系统管理的存储空间中。
7. 如权利要求 1-6 任一项所述的系统，其特征在于，
- 所述优化者服务模块还用于：鉴定所述模型优化者是否具有调用所述预训练模型对应的已混淆模型参数的权限；若所述模型优化者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型并提供至所述模型优化者；和/或，
- 35 所述消费者服务模块还用于：鉴定所述模型消费者是否具有调用所述预训练模型或所述优化后模型对应的已混淆模型参数的权限；若所述模型消费者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以

获取所述预训练模型或所述优化后模型并提供至所述模型消费者。

8. 如权利要求 1-7 任一项所述的系统，其特征在于，所述预训练服务系统针对所述预训练模型或所述优化后模型维护有独立的模型存取代码和模型训练推理代码；其中，所述模型结构代码以及所述模型存取代码由所述预训练服务系统经加密后维护，

5 所述模型训练推理代码由所述预训练服务系统明文维护。

9. 一种基于预训练服务系统的服务提供方法，其特征在于，所述预训练服务系统包括生产者服务模块、优化者服务模块和消费者服务模块；所述预训练服务系统所维护的模型包括已混淆模型结构代码和已混淆模型参数，所述已混淆模型结构代码与所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理

10 后生成；所述方法包括：

获取访问所述预训练服务系统的目标用户的身份信息，以确定所述目标用户所属的预设用户类型；

向所述目标用户开放与其所属的预设用户类型相匹配的服务模块，其中：

在所述目标用户为模型生产者的情况下，向所述目标用户开放所述生产者服务模块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；

15 块，所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能，并产出相应的预训练模型；

在所述目标用户为模型优化者的情况下，向所述目标用户开放所述优化者服务模块，所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型进行优化，并得到优化后模型；

在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

20 在所述目标用户为模型消费者的情况下，向所述目标用户开放所述消费者服务模块，所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口，所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

10. 如权利要求 9 所述的方法，其特征在于，所述模型预训练功能用于根据所述预训练数据集训练得到通用预训练模型；其中，所述预训练模型包括以下至少之一：所述通用预训练模型、所述生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。

25 预训练数据集训练得到通用预训练模型；其中，所述预训练模型包括以下至少之一：所述通用预训练模型、所述生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。

11. 如权利要求 10 所述的方法，其特征在于，还包括：

生成所述专用预训练模型对应的代码开发模板；

30 将所述代码开发模板提供至所述模型优化者；以及，根据所述模型优化者发出的模型修改指令，对所述专用预训练模型的模型结构代码和/或模型参数进行修改。

12. 如权利要求 9-11 任一项所述的方法，其特征在于，还包括：

通过所述服务接口将所述预训练模型或所述优化后模型部署于所述模型消费者对应的设备，以用于对所述模型消费者提供的数据进行离线推理；或者，

35 通过所述服务接口获取所述模型消费者提供的数据，并调用所述预训练模型或所述优化后模型进行在线推理。

13. 如权利要求 9-12 任一项所述的方法，其特征在于，所述相关联的混淆处理包括：对所述原始模型结构代码中的至少一部分代码进行正向混淆处理，并对所述原始模型参数中与所述至少一部分代码相关的模型参数进行逆向混淆处理。

14. 如权利要求 9-13 任一项所述的方法, 其特征在于, 所述已混淆模型参数存储在所述预训练服务系统管理的存储空间中。

15. 如权利要求 9-14 任一项所述的方法, 其特征在于, 还包括:

5 鉴定所述模型优化者是否具有调用所述预训练模型对应的已混淆模型参数的权限; 若所述模型优化者具有调用所述已混淆模型参数的权限, 则将所述已混淆模型参数与
所述已混淆模型结构代码相结合以获取所述预训练模型并提供至所述模型优化者; 和/或,

10 鉴定所述模型消费者是否具有调用所述预训练模型或所述优化后模型对应的已混淆模型参数的权限; 若所述模型消费者具有调用所述已混淆模型参数的权限, 则将所述已混淆模型参数与
所述已混淆模型结构代码相结合以获取所述预训练模型或所述优化后模型并提供至所述模型消费者。

16. 如权利要求 9-15 任一项所述的方法, 其特征在于, 所述预训练服务系统针对所述目标预训练模型维护有独立的模型存取代码和模型训练推理代码; 其中, 所述模型
15 结构代码以及所述模型存取代码由所述预训练服务系统经加密后维护, 所述模型训练推理代码由所述预训练服务系统明文维护。

17. 一种基于预训练服务系统的服务提供装置, 其特征在于, 所述预训练服务系统包括生产者服务模块、优化者服务模块和消费者服务模块; 所述预训练服务系统所
维护的模型包括已混淆模型结构代码和已混淆模型参数, 所述已混淆模型结构代码与
20 所述已混淆模型参数为预先针对原始模型结构代码和原始模型参数进行相关联的混淆处理后生成; 所述装置包括:

获取单元, 用于获取访问所述预训练服务系统的目标用户的身份信息, 以确定所述目标用户所属的预设用户类型;

开放单元, 用于向所述目标用户开放与其所属的预设用户类型相匹配的服务模块, 其中:

25 在所述目标用户为模型生产者的情况下, 向所述目标用户开放所述生产者服务模块, 所述生产者服务模块用于向所述模型生产者提供针对预训练数据集的模型预训练功能, 并产出相应的预训练模型;

在所述目标用户为模型优化者的情况下, 向所述目标用户开放所述优化者服务模块, 所述优化者服务模块用于根据所述模型优化者提供的微调数据集对所述预训练模型
30 进行优化, 并得到优化后模型;

在所述目标用户为模型消费者的情况下, 向所述目标用户开放所述消费者服务模块, 所述消费者服务模块用于向所述模型消费者提供针对所述预训练模型和/或所述优化后模型的服务接口, 所述预训练模型或所述优化后模型用于对所述模型消费者提供的数据进行推理并输出模型预测结果。

35 18. 如权利要求 17 所述的装置, 其特征在于, 所述模型预训练功能用于根据所述预训练数据集训练得到通用预训练模型; 其中, 所述预训练模型包括以下至少之一: 所述通用预训练模型、所述生产者服务模块根据所述通用预训练模型生成的对应于所述模型生产者所指示的下游任务的专用预训练模型。

19. 如权利要求 18 所述的装置, 其特征在于, 所述装置还包括:

生成单元，用于生成所述专用预训练模型对应的代码开发模板；

提供及修改单元，用于将所述代码开发模板提供至所述模型优化者；以及，根据所述模型优化者发出的模型修改指令，对所述专用预训练模型的模型结构代码和/或模型参数进行修改。

5 20. 如权利要求 17-19 任一项所述的装置，其特征在于，所述装置还包括：

离线部署单元，用于通过所述服务接口将所述预训练模型或所述优化后模型部署于所述模型消费者对应的设备，以用于对所述模型消费者提供的数据进行离线推理；或者，

10 在线推理单元，用于通过所述服务接口获取所述模型消费者提供的数据，并调用所述预训练模型或所述优化后模型进行在线推理。

21. 如权利要求 17-20 任一项所述的装置，其特征在于，所述相关联的混淆处理包括：对所述原始模型结构代码中的至少一部分代码进行正向混淆处理，并对所述原始模型参数中与所述至少一部分代码相关的模型参数进行逆向混淆处理。

15 22. 如权利要求 17-21 任一项所述的装置，其特征在于，所述已混淆模型参数存储在所述预训练服务系统管理的存储空间中。

23. 如权利要求 17-22 任一项所述的装置，其特征在于，所述装置还包括：

20 第一鉴定单元，用于鉴定所述模型优化者是否具有调用所述预训练模型对应的已混淆模型参数的权限；若所述模型优化者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型并提供至所述模型优化者；和/或，

第二鉴定单元，用于鉴定所述模型消费者是否具有调用所述预训练模型或所述优化后模型对应的已混淆模型参数的权限；若所述模型消费者具有调用所述已混淆模型参数的权限，则将所述已混淆模型参数与所述已混淆模型结构代码相结合以获取所述预训练模型或所述优化后模型并提供至所述模型消费者。

25 24. 如权利要求 17-23 任一项所述的装置，其特征在于，所述预训练服务系统针对所述目标预训练模型维护有独立的模型存取代码和模型训练推理代码；其中，所述模型结构代码以及所述模型存取代码由所述预训练服务系统经加密后维护，所述模型训练推理代码由所述预训练服务系统明文维护。

25. 一种电子设备，其特征在于，包括：处理器、内存及非易失性存储器；

30 所述处理器用于从所述非易失性存储器中读取对应的计算机程序到所述内存中，并基于所述计算机程序执行如权利要求 9-16 任一项所述的方法。

26. 一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求 9-16 任一项所述的方法。

35 27. 一种计算机程序产品，其特征在于，包括计算机程序，当所述计算机程序被运行时，使得计算机执行如权利要求 9-16 任一项所述的方法。