《PerFED-GAN: Personalized Federated Learning via Generative Adversarial Networks》

联邦学习作为一种分布式机器学习方法越来越受欢迎，可用于部署依赖人工智能的物联网应用程序，同时保护客户端数据隐私和安全。由于客户的差异，单一的全局模型可能无法在所有客户上都表现良好，因此为每个客户训练出更适合其个人需求的个性化模型的个性化联邦学习方法成为研究热点。然而，大多数个性化的联邦学习研究都侧重于数据异质性，而忽略了模型架构异质性的需求。现有的联邦学习方法大多统一设置所有参与联邦学习的客户端的模型架构，不方便每个客户端的个体模型和本地数据分布需求，也增加了客户端模型泄露的风险。本文提出了一种基于协同训练和生成对抗网络（GAN）的联邦学习方法，该方法允许每个客户端设计自己的模型以独立参与联邦学习训练，而无需与其他客户端或中心共享任何模型架构或参数信息。在我们的实验中，当客户端的模型架构和数据分布发生显着变化时，所提出的方法在平均测试准确率方面优于现有方法 42%。

联邦学习的有效性源于客户之间的知识共享，而个性化联邦学习旨在在客户之间共享共同知识，同时保持他们的个性化知识。 现有的联邦学习方法通过模型参数共享知识，需要一致的模型架构，这不仅削弱了客户端模型的个性，而且增加了模型泄漏的风险。

相关工作

1、Personalized Federated Learning

最近提出许多个性化方案，这些方法都需要客户端上传其模型参数以进行全局聚合，这可能会泄漏客户端模型。最近，已经引入了几种安全计算技术来保护联邦学习中的数据和模型参数，包括差分隐私、安全多方计算、同态加密和可信执行环境，但是这些方法仍然存在一些缺点，例如大量的通信或计算成本，或者依赖特定的硬件来实现。

这些用于个性化任务的联邦学习方法的另一个限制是每个客户模型的架构必须一致，因为模型参数通常在这些方法的联邦学习训练过程中聚合或对齐。它阻止客户独立设计独特的模型架构。

FedMD通过在公共数据集上对齐多个神经网络的 logits，利用知识蒸馏来传达每个客户本地数据的知识。FedMD 的客户端模型可以是各种架构的神经网络，只要它们在 logits 输出层中保持一致即可。然而，FedMD 需要大量的标记数据作为客户端模型对齐的公共数据集。标记数据采集难度大，开放的大规模标记数据集通常很少，这限制了 FedMD 的应用场景。

Guha 等人使用蒸馏模型生成全局集成模型，该模型需要共享客户的本地模型或他们的蒸馏模型。但是，它将客户端的私有数据或本地模型信息暴露给中央服务器，这使得该解决方案不适合希望避免模型泄漏的客户。

2、Co-training

在两个视图上训练两个不同的分类器。之后，两个分类器都为另一个分类器伪标记一些未标记的实例。然后每个分类器可以在其原始训练集上重新训练，新的训练集被另一个分类器伪标记。
Wang 等人证明，如果两个视图具有较大的多样性，则协同训练几乎不会受到标签噪声和采样偏差的影响，并且即使视图不足，也可以通过利用未标记数据来输出最佳分类器的近似值。
对于单视图设置，Zhou 等人提出了三重训练，它使用 3 个具有较大多样性的分类器来投票选出最自信的未标记数据的伪标签。 此外，王等人证明具有较大的多样性的分类器也可以在单一视图设置中提高模型性能。同时，他们还指出了分类器在协同训练中的性能在几轮后饱和的原因。 即分类器相互学习并逐渐变得相似，它们的差异不足以支持进一步的性能提升。

个性化联合学习设置中的不同客户端通常没有足够的本地数据，这些数据仅是总体分布的一方面，尤其是在非 IID 数据设置中。 因此，根据来自多个客户的本地数据训练的模型可能具有高度的多样性，这可能导致协同训练发挥作用。

3、Generative Adversarial Network

生成对抗网络（GAN）是 Ian Goodfellow 和他的同事在 2014 年设计的一种机器学习框架。对于给定的训练集，该技术学习生成具有相同分布的新数据样本。

GAN 主要由两部分组成：判别器网络和生成器网络。GAN的核心思想是通过判别器网络间接训练，判别器网络本身也是动态更新的。这基本上意味着生成器网络没有经过训练来最小化与特定对象的距离，而是用于欺骗判别器网络。

生成网络生成候选者，判别器网络评估它们。竞争在数据分布方面进行。通常，生成网络学习从潜在空间映射到感兴趣的数据分布，判别器网络将生成器生成的候选者与真实数据分布区分开来。生成器网络的训练目标是提高判别器网络的错误率。

初始训练数据集用于判别器网络。训练过程涉及向其展示训练数据集中的样本。根据生成器是否成功欺骗判别器来训练生成器。通常，生成器的种子是从预定义的潜在空间中采样的随机输入。此后，由生成器合成的候选者由判别器评估。

问题陈述

考虑到个性化的联邦学习设置，有 N 个客户端，每个客户端 $C_i(1 \le i \le N)$ 都有来自分布 D_i 的私有数据集 D_i。 由于个性化联邦学习场景中的不同客户端数据一般都是Non-IID，即：

$$\exists i \ne j(1 \le i, j \le N) \to D_i \ne D_j \quad (1)$$

每个 C_i 也有其模型 $M_i(w_i)$，其中 M_i 是模型架构，w_i 是 M_i 的相应模型参数。在兼容异构模型架构的个性化联邦学习场景中，我们假设不同客户端的模型架构可以不同，即：

$$\exists i \ne j(1 \le i, j \le N) \to M_i \ne M_j \quad (2)$$

这时，M_i 的 w_i 和 M_j 的 w_j 也不同。差异不仅是参数值的差异，甚至可能是参数的数量，因为不同的模型架构可能有不同的模型参数数量。即使在 M_i 和 M_j 的参数个数相等的巧合情况下，w_i 和 w_j 的参数值也有不同的含义，因此比较它们的值是没有意义的。

定义一个函数 $f(x; M_i, w_i)$ 来评估模型 $M_i(w_i)$ 对样本 x 的性能，f 函数的值越小意味着模型对 x 的性能越好。对于分类任务，当 x 被 $M_i(w_i)$ 正确分类时，f 的输出为 0，否则为 1。对于客户端 c_i，其任务目标是针对其给定的模型架构 M_i 优化 w_i，以最小化函数 f 对 D_i 的期望：

$$w_i^* = \underset{w_i}{arg\ min}\ \underset{x \sim D_i}{E}[f(x; M_i, w_i)] \quad (3)$$

$E(\cdot)$ 是期望函数。对于分类任务，公式 (3) 表示优化 w_i，以最小化模型 M_i 的泛化分类误差。假设客户端 C_i 本地训练得到的最优模型参数为 $w_{i,loc}^*$ 对于 $i = 1,2,...,N$：

$$w_{i,loc}^* \overset{local}{=} \underset{w_i}{arg\ min}\ \underset{x \sim D_i}{E}[f(x; M_i, w_i)] \quad (4)$$

个性化联邦学习的目的是为每个客户端协作训练个性化模型 $M_i(w_i)$，而不将 D_i 暴露给 $C_j(1 \le i, j \le N, i \ne j)$：

$$w_{i,fed}^* \overset{federated}{=} \underset{w_1, w_2, ..., w_N}{arg\ min}\ \frac{1}{N}\sum_{i=1}^{N} \underset{x \sim D_i}{E}[f(x; M_i, w_i)] \quad (5)$$

联邦训练得到的模型的性能应该足够高。具体来说，联邦学习中每个客户端的模型性能不应低于 $i = 1,2,...,N$ 的本地训练模型性能：

$$\underset{x \sim D_i}{E}[f(x; M_i, w_{i,fed}^*)] \le \underset{x \sim D_i}{E}[f(x; M_i, w_{i,loc}^*)] \quad (6)$$

对于分类任务，即每个参与者的联合训练模型的分类泛化精度不低于本地训练模型。

PerFED-GAN 的总体步骤

联邦学习的主要动机是提高客户端模型的性能。本地训练模型性能不佳的原因通常是缺乏本地数据，这阻止了模型通过仅在本地数据集上进行训练来学习足够的任务专业知识。因此，为了提高客户端模型的性能，它们必须能够向其他客户端学习。最流行和最直接的知识交流方式是共享数据或模型，但是这些在我们的联邦学习环境中是不允许的，因此我们需要开发替代方法。

研究相关的协同训练表明，为了提高分类任务模型的性能，模型之间需要足够大的多样性。一般来说，在单一视图设置中生成具有显着差异的模型是很困难的。然而，在联邦学习环境中，客户端模型可能具有完全不同的架构，并在很可能是非 IID 的个性化私有数据集上进行训练。所有这些因素都可能导致不同客户的模型发生变化。 因此，我们的中心思想是使用在每个客户端本地数据上训练的模型作为判别器网络，并生成新的样本。在收到这些创建的数据集后，其他客户使用它们来进一步训练他们的本地模型，从而提高他们定制模型的性能。PerFED-GAN 的总体步骤如下。

1、本地模型训练：每个客户端在其本地数据集上独立训练其模型。

2、GAN 训练：每个客户端使用上一步训练的本地模型作为判别器网络来训练生成器网络，并用它来生成新的数据集。

3、通信：中心根据标签汇总从客户端采集的生成数据样本，并将结果发送回各个客户端。

4、客户端模型更新：每个客户端通过在从其本地数据集和接收到的新数据集合并的新数据集上对其进行训练来更新其模型。之后，循环执行步骤 2 到步骤 4 ，进一步提高个性化联邦学习模型的性能。

in-process training是在其私有数据集和从中心接收的聚合数据集上训练客户端模型。每个客户端从中心接收到的聚合数据集包含从其他客户端生成的样本。这些样本及其对应的标签可以看作是其他客户端对未标记样本的分类结果，因为这些样本是其他客户端根据目标标签生成的。此时，使用聚合数据集进行in-process training相当于使用协同训练中使用另一个分类器标记未标记数据的结果进行训练。

PerFED-GAN 中使用的本地模型训练方法和 GAN 训练方法是模块化和可替换的。 理论上，几乎所有的神经网络训练方法和 GAN 训练方法都可以应用于 PerFED-GAN。

在 GAN 训练阶段之后，中央服务器从每个客户端在聚合过程中生成的数据中收集所有类别的数据。例如，有3个类别，A，B，C。客户端 1 生成数据集{A1，A2，B1，B2，C1，C2}，客户端 2 生成数据集{A3，A4，B3，B4，C3， C4}，客户端 3 生成数据集 {A5, A6, B5, B6, C5, C6}。中心服务器将这些聚合到一个大数据集{A1，A2，...，A6，B1，B2，...，B6，C1，C2，...，C6}，然后，中心随机选择一些样本从大型数据集中发送到每个客户端。

B、算法分析

PerFED-GAN 是一种个性化的联邦学习方案，其特点是支持不同模型架构的客户端，其详细算法流程如算法 1 所示。M_i 为模型架构，w_i 为 M_i 对应的模型参数。GM_i 是带有参数 v_i 的生成器网络。Tr_i 是客户端 C_i 的本地训练数据集。MAXROUND 是最大通信轮数。A_i 是包含 GM_i 生成的样本及其生成标签的数据集。Ca 是在中心服务器中分配的一个数据集，用于存储每个客户端上传的 A_i 中的样本。Tr_i' 是一个数据集，包含 Tr_i' 中的所有样本和来自 Ca 的随机选择的样本。

PerFED-GAN 算法的某些部分是可独立替换的，包括本地模型初始化训练方法、GAN 训练方法和本地模型更新训练方法。理论上，任何可以用来训练神经网络的方法都可以作为替代。因此，PerFED-GAN 可以利用最新的神经网络优化研究成果来增强训练效果，几乎不需要做更多额外的改变。

算法的客户端执行部分可以由不同的客户端并行异步执行，在中心服务器执行的任务中，合并各个客户端上传的数据集的时间复杂度主要与样本总数成正比由所有客户上传。上传样本的数量可以根据实际情况进行调整，在通信带宽足够的情况下可以上传更多的样本。此外，将生成网络直接上传到中心，由中心生成样本也是一种可行的策略，但也可能带来模型隐私泄露的额外风险。此外，上传生成模型或生成样本对训练时间效率的影响还应根据通信能力、样本量、生成模型的大小、中心服务器的计算资源等各种因素进行评估。

C、收敛分析

D、超参数

在定理 1 中，公式 (11) 的左边随着 $M = ga_0$ 而增加，这意味着，对于一个固定的 a_0 ，一个给定的 h_2 其泛化误差为 b_0 和大小为 s_2 的 D_2 ，数据集 G 的大小足够大 需要 h_1 。

在 PerFED-GAN 的实际应用中，每个客户端的私有训练数据集的大小，即 s_2 可能不同。由式 (11) 可以看出，当 s_2 较大时，需要生成的数据集较大才能满足式 (11) 的要求。不等式成立，因此我们使用超参数 β 来确定所需生成数据集的大小，即

$$g = \beta s_2 \quad (23)$$

这样，可以设置一个合理的超参数 β ，为不同私有数据集的客户端提供足够数量的生成样本，同时避免过度的样本生成要求带来的高通信成本。

E、具有差分隐私的 PerFED-GAN

GAN 可能会生成与原始学习样本相似的样本，从而导致上传到 PerFED-GAN 的生成数据存在数据隐私泄露的可能性。在实验部分，我们展示了可以通过限制 GAN 训练轮数来降低生成样本的质量，以降低泄露原始数据的风险，实验结果表明，减少 GAN 训练轮数会导致模型质量稍微下降。此外，保护原始数据隐私的进一步措施是引入差分隐私 GAN。Torkzadehmahani 等人提出了一种差分隐私条件GAN，DP-CGAN，可用于生成指定类别的数据并保留训练数据的隐私。因此，PerFED-GAN 中的 GAN 训练模块可以替换为 DP-CGAN 或具有类似 DP 技术的 GAN 训练算法，可以有效避免生成样本泄露训练数据隐私。

6、结论

在本文中，我们提出了与异构模型架构兼容的 PerFED-GAN 联邦学习方法。PerFED-GAN 比现有方法更适合异质性和个性化需求更高的个性化联邦学习设置。此外，PerFED-GAN 不仅保护联邦学习环境中所有客户的私人数据，还保护他们的私人模型和训练策略。PerFED-GAN 使客户能够共享多方知识，以提高其本地模型的性能。它在具有不同架构的异构模型的非 IID 数据集上产生了有希望的结果，这些模型更实用，但在现有的联邦学习方法中通常难以处理。