



计算机科学

COMPUTER SCIENCE

聚类联邦学习簇间优化

李仁杰, 闫巧

引用本文

李仁杰, 闫巧. 聚类联邦学习簇间优化[J]. 计算机科学, 2023, 50(11A): 221000243-5.

LI Renjie, YAN Qiao. [Inter-cluster Optimization for Cluster Federated Learning](#)[J]. Computer Science, 2023, 50(11A): 221000243-5.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[二维多边形物体剖分研究](#)

Study on Decomposition of Two-dimensional Polygonal Objects

计算机科学, 2023, 50(11A): 230300237-5. <https://doi.org/10.11896/jsjcx.230300237>

[一种面向多模态医疗数据的联邦学习隐私保护方法](#)

Federated Learning Privacy-preserving Approach for Multimodal Medical Data

计算机科学, 2023, 50(11A): 230800021-8. <https://doi.org/10.11896/jsjcx.230800021>

[一种基于CutMix的增强联邦学习框架](#)

Enhanced Federated Learning Frameworks Based on CutMix

计算机科学, 2023, 50(11A): 220800021-8. <https://doi.org/10.11896/jsjcx.220800021>

[基于配置语句树的网络设备配置异常检测算法](#)

Anomaly Detection Algorithm for Network Device Configuration Based on Configuration Statement Tree

计算机科学, 2023, 50(11A): 230200128-10. <https://doi.org/10.11896/jsjcx.230200128>

[基于物品关联协同过滤的下一购物篮推荐算法](#)

Next-basket Recommendation Algorithm Based on Correlation Between Items Collaborative Filtering

计算机科学, 2023, 50(11A): 221000076-6. <https://doi.org/10.11896/jsjcx.221000076>

聚类联邦学习簇间优化

李仁杰 闫巧

深圳大学计算机与软件学院 广东 深圳 518000

(1441347985@qq.com)

摘要 聚类联邦学习常用于解决联邦学习中数据异质性导致准确率下降的问题,通过聚类算法将数据分布相似的客户端划分到相同簇中,簇模型用于解决某个特定分布问题。当前聚类联邦学习中为了获得好的实验结果,研究者通常将相同的分布训练集和测试集分配到同一个簇中,然而现实中无法达到实验理想效果,本地客户端中使用模型的数据集与训练模型的数据集分布可能不同,当分布不同时聚类联邦学习的簇模型准确率会大幅下降,影响本地端设备的簇间准确率。文中提出两种方案提升聚类联邦学习中簇模型的簇间准确率。第一种方案是自适应聚类联邦学习(AWCFL),在簇内聚合时加入其他簇的模型,使得簇模型学习到其他分布的知识,有效提升簇模型的簇间准确率;第二种方案是多分布聚类联邦学习(MCFL),将簇模型同步到每个客户端,让客户端选择合适的模型使用,该方案相对于第一种方案簇内准确率不会下降,簇间准确率提升明显。上述两种方案在Mnist和EMnist数据集上进行实验,与IFCA,CFL(Clustered Federated Learning)和FedAvg进行比较,簇间准确率明显提升。

关键词: 联邦学习;聚类;聚合优化;数据分布

中图法分类号 TP301

Inter-cluster Optimization for Cluster Federated Learning

LI Renjie and YAN Qiao

School of Computer and Software, Shenzhen University, Shenzhen, Guangdong 518000, China

Abstract Clustering federated learning is often used to solve the problem of decreasing accuracy caused by data heterogeneity in federated learning. The idea is to group clients with similar data distributions into the same cluster using clustering algorithms, and then train a cluster model specifically for that distribution. However, in practical applications, it is challenging to achieve ideal results because the training and test datasets used by the local clients may not match the data distribution of the cluster model, leading to a significant drop in inter-cluster accuracy. To improve the accuracy of the cluster model in clustering federated learning, this paper proposes two solutions. The first is adaptive weighted clustering federated learning(AWCFL), which incorporates models from other clusters during intra-cluster aggregation, enabling the cluster model to learn from other distributions and effectively improve inter-cluster accuracy. The second solution is multi-distribution clustering federated learning(MCFL), which synchronizes the cluster model with each client, allowing clients to choose the appropriate model to use. Compared with the first solution, intra-cluster accuracy remains unaffected in MCFL, while inter-cluster accuracy is significantly improved. To evaluate the proposed solutions, experiments are conducted on the Mnist and EMnist datasets. Compared with IFCA, clustered federated learning(CFL) and FedAvg, the accuracy rate between clusters is significantly improved.

Keywords Federated learning, Clustering, Aggregation optimization, Data distribution

1 引言

随着互联网的发展,移动设备获取使用者的数据进行训练,能极大提升设备使用者的用户体验,然而,需要个人隐私数据引发了人们对个人隐私泄露的担忧从而拒绝上传数据。为了解决个人隐私与用户体验的冲突,联邦学习应运而生,McMahan等^[1]于2016年提出了一种创新的隐私保护机器学习范式,它在参与的用户设备上分发协作模型训练,而不需要用户共享他们的原始训练样本。在实践中,联邦学习服务器

首先将全局模型分发给随机参与训练的客户端。然后,每个客户端根据自己的局部数据并行使用梯度下降算法优化自己的局部模型。最后,联邦学习服务器对所有局部模型的更新或参数进行平均,并将其聚合起来构建一个新的全局模型。

Hard等^[2]使用联邦学习在诸如下一个单词预测等应用中和Brisimi等^[3]学习电子健康记录都取得了成功。但是大多数情况下,客户的本地数据是非独立同分布的,导致联邦学习出现模型训练缓慢、准确率较差等问题。为了应对这一挑战,最近的一些工作提出通过执行局部模型个性化来绕过

基金项目:国家自然科学基金(61976142);深圳市科技计划项目(JCYJ20210324093609025)

This work was supported by the National Natural Science Foundation of China(61976142) and Shenzhen Science and Technology Plan Project(JCYJ20210324093609025).

通信作者:闫巧(yanq@szu.edu.cn)

数据异构性。这些算法的训练目标不是追求一个所有客户共享的普遍适用的模型,而是为每个客户创建一个适合其本地数据的模型。个性化方法包括局部微调^[4]、模型内插^[5]。本文聚焦于另一种方法:聚类联邦学习。

聚类联邦学习放宽了联邦学习的限制,即它不要求每个客户端具有相同的数据分布而是允许不同的客户端存在不同数据分布。聚类联邦学习的目标是为每个分布训练一个模型。在传统的聚类联邦学习中,一个客户端只能与一个数据分布相关联。在聚类联邦学习中,non-IID(非独立同分布)问题很容易解决:只需将具有相同数据分布的客户端分组到一个簇中,然后在每个簇上进行常规联邦学习,此时数据分布在簇中是 IID(独立同分布)^[6]。

然而,聚类联邦学习在实践中面临一个基本问题:在部署聚类联邦学习时,难以对客户使用者进行判定,不太可能拥有相同的数据分布。事实上,现实世界的用户数据更有可能遵循多种混合分布^[7]。

上述问题主要是由于缺乏簇间交互导致每个簇内的知识学习数据分布单一,同时每个本地端使用簇模型时无法确保数据分布相同,导致数据出现不同分布时簇间准确率会大幅下降,无法给予用户良好的体验。为了解决上述问题,本文提出两种优化方案以提高簇模型的泛用性,通过对所提出的优化方案进行在不同数据集下的对比实验,验证了所提出的两种方案都能够提高聚类联邦学习的簇间准确率。本文主要的创新点如下:

1) 提出自适应权重聚类联邦学习 AWCFL (Adaptive Weighted Clustering Federated Learning) 方案,在簇聚合时将其他簇以低权重加入聚合,让本地端模型能够学习簇外的数据分布的知识,通过修改聚合方式让不同簇之间进行梯度交互,提升簇模型的泛用性。

2) 提出多分布聚类联邦学习 MCFL (Multi-distribution Cluster Federated Learning) 方案,将所有簇模型加载到每个客户端,客户端可以根据数据分布选择合适的模型使用,该方案能够有效提升簇间准确率。

3) 在对比实验中,我们将所提出两种优化方案 AWCFL, MCFL 和传统的模型 CFL, FedAvg 进行比较。实验结果表明,两种优化方案都能够有效提升聚类联邦学习的簇间准确率。

2 相关工作

2.1 联邦学习

Mcmahan 等首先引入了联邦学习^[1]和普通联邦学习优化方法 FedAvg^[7],旨在为分布式机器学习模型训练提供隐私保护支持。

如图 1 所示,联邦学习系统^[8]包括一个维护全局模型的云服务器和多个参与设备,这些设备通过网络与服务器通信。在式(1)中每一轮通信 t 中,服务器从客户端中随机抽取一个子集 K 参加这一轮的训练。然后,服务器将最新的全局模型 w_t 广播给这些客户端,以进行进一步的局部优化。然后,每个参与的客户端通过局部数据集以多个局部 epoch 并行优化设备数据的局部目标函数。然后,对这些在时间预算内完成培训的客户端本地计算的参数更新 w_t 进行汇总。最后,服务器将全局模型更新为 w_{t+1} ,完成本轮学习。

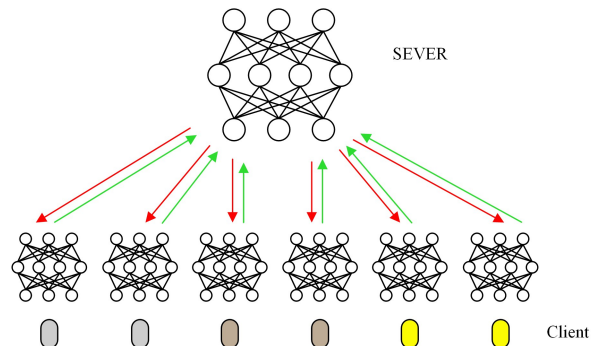


图 1 联邦学习

Fig. 1 Federated learning

$$W_{i+1}^*(x) := W_{i+1}^*(x) - \eta \left(\frac{\sum_i \Delta w_i}{n} \right) \quad (1)$$

2.2 聚类联邦学习

图 2 为聚类联邦学习模型图,聚类联邦学习是按照某个特定标准(如距离)把若干个客户端分到不同的簇中,使得同一个簇内的本地模型趋同,不在同一个簇中的模型的差异性也尽可能大。簇内相似模型融合不会导致模型相互影响。

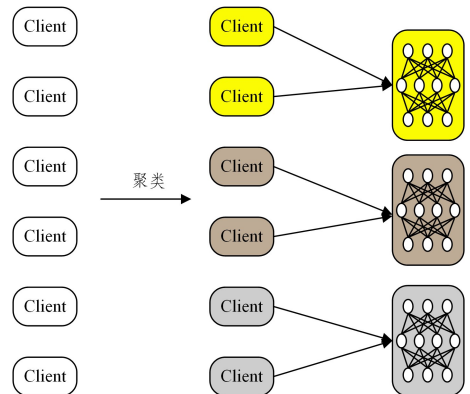


图 2 聚类联邦学习

Fig. 2 Clustering federated learning

如前所述,联邦学习的主要挑战之一是统计异质性(如数据非独立同分布^[8]、大小不平衡^[9]、类别不平衡^[10])。传统的方法是在不一致的数据上训练一个一致的全局模型,这在高异质性设置^[11]中会导致性能不佳。CFL 将优化目标划分为几个子目标,并遵循多元组架构^[12],而不是对一个共识全局模型进行优化。CFL 维护多个组(或集群)模型,这些模型比共识模型更专业,并实现了较高的准确性。

Sattler 等提出了第一个基于聚类联邦学习(CFL)的框架^[13],递归分离下降方向不一致的两组客户端。作者进一步提出^[14]来提高基于聚类联邦学习的框架在拜占庭环境中的鲁棒性。但 CFL 采用的递归二分法计算效率低,需要多轮通信才能将所有不一致的客户端完全分离。此外,由于每一轮参与的客户端是随机的,这可能会导致递归分区算法失败。为了提高 CFL 的效率,Ghosh 等提出了 IFCA^[9],该算法随机生成簇中心,并将客户端划分为最小化损失值的簇。虽然 IFCA 的模型精度显著提高,但 IFCA 需要在每一轮广播所有组模型,并且对初始化具有概率成功的敏感性。FeSEM^[14]使用基于距离的随机期望最大化(EM)算法,而不是基于距离的邻域方法。然而,距离在高维、低样本量(HDLSS)情况下经常发生,即高维^[15]中的距离集中现象。距离浓度会导致邻域

结构^[16]被破坏,这会对 K-Means, k -Medoids 和层次聚类等基于距离的两两聚类算法的性能产生不利影响。类似于基于余弦相似度的聚类方法^[17], Briggs 等提出了一种凝聚层次聚类方法^[15],它依赖于迭代计算所有簇之间的两两距离,计算量较大。值得注意的是,上述所有 CFL 框架都假设所有客户端都将参与集群过程,没有新设备,这在大型联邦训练系统中是不现实的。

3 聚类联邦学习簇间优化

在传统的聚类联邦学习中,将客户端分类后每个簇都将形成一个信息孤岛,簇内的客户端通过簇服务器学习簇内所含数据集的知识。因此每个簇模型针对某一类分布都有良好的效果,相反,遇到不合适的分布数据,准确率会严重下滑。在实验中研究者通常自主选取合适的簇模型以达到良好的效果,但是在现实生活中无法确认模型使用者适合哪类分布的簇模型,使用者采用不合适的簇模型,准确率将大幅下降。本文为了解决上述问题,提出两种优化方案来兼容簇模型间的差异,提升簇模型的泛用性,使得簇模型针对不同分布的数据都有良好的训练效果。为了衡量提升效果,实验中采用新的评价指标——簇间准确率(簇模型对其他分布数据的识别准确率)。

3.1 AWCFL

为了提高簇内模型对多分布的数据的准确率,本文提出第一种优化方案:在簇聚合时本簇和其他簇的神经网络模型共同聚合。图 3 即 AWCFL 优化后的模型图。

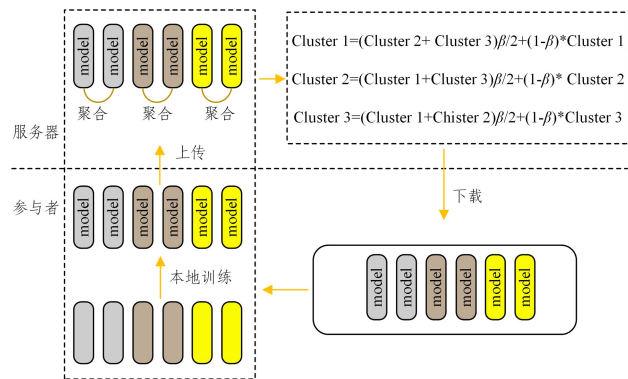


图 3 AWCFL

Fig. 3 AWCFL

AWCFL 优化方案为了让优化后的聚类联邦学习能够兼顾簇间和簇内的学习效率,聚合时采用不同的权重,聚合公式如下:

$$C_k(x) = \frac{\beta}{n-1} \sum_{i=1, i \neq k}^n C_i(x) + (1-\beta) C_k(x) \quad (2)$$

其中, C 代表簇模型; $\beta/(n-1)$ 是其他簇加入到聚合的权重, $(1-\beta)$ 是本簇的权重更新。 β 的变化会影响聚合程度, β 越大,聚合程度越高,簇内准确率下降,簇间准确率提升; β 越小,聚合程度越低,簇内准确率提升,簇间准确率降低;当 $\beta=0$ 时,簇间不聚合,也就是传统聚类联邦学习。经过实验发现,当 $\beta=0.5$ 时,簇间准确率和簇内准确率达到合适的权重。第一种优化方案具体算法如算法 1 所示。

算法 1 AWCFL

输入:全局模型 W^* ,簇模型 Cluster,簇的数量 K ,参与训练的客户端数量 N ,客户端模型 w_i

输出:全局模型, K 个簇模型, N 个客户端模型

1. /* AWCFL 的流程 */
2. $w_1, w_2, w_3, \dots, w_n \leftarrow W^*$
3. for $t \leftarrow 1$ to T do
 - 3.1. $w_n^i = w_n^{i-1} - \eta \nabla \text{Loss}(w_n^{i-1}, x_n), \forall n \in \{1, 2, 3, \dots, N\}$
 - 3.2. $c_1, c_2, c_3, \dots, c_K \leftarrow K\text{-means}(w_1, w_2, \dots, w_N) / * \Gamma$ 为簇的集合 */
 - 3.3. for $k \leftarrow 1$ to K do
 - 3.3.1. $c_k(x) = \frac{\beta}{n-1} \sum_{i=1, i \neq k}^n c_i(x) + (1-\beta) \frac{1}{\text{len}(c_k)} \sum_{i=1}^{\text{len}(c_k)} w_i(x)$
 - 3.4. $w_1, w_2, w_3, \dots, w_n \leftarrow c_1, c_2, c_3, \dots, c_K$
4. return($W^*, w_1, w_2, \dots, w_N$)

3.2 MCFL

AWCFL 方案可以有效提高簇间准确率,然而,簇间聚合采用不同簇时,不同分布的混合会导致簇内学习率下降。为了解决这个问题,本文提出了第二种优化方案,能在不影响原有簇内准确率的前提下提高簇间准确率。如算法 2 所示,客户端存储每个簇模型,即存储每个分布的簇模型,在应用模型时,参与者选择适配的模型使用。

最适合测试集模型应该是与测试集分布类似的簇模型,MCFL 采用的方法是利用特征均值,特征均值能够有效显示数据分布情况,每个参与者依据自身数据集采用 n 个样本图片,计算均值 \bar{X}_i ,上传到服务器中,每个簇中心对簇内参与者上传的 \bar{X}_i 进行平均,并作为簇模型的特征 \bar{X} ,测试集依据数据采用 n 个数据并计算均值得到 μ ,利用 $t = |\bar{X} - \mu|$ 基于特征均值判断测试集数据分布与哪一个簇模型分布接近,选择 t 最小的簇模型使用,也就是找到分布相近的簇模型。

图 4 为采用 MCFL 优化后的模型架构图,在使用测试集时,利用测试集生成自己的特征图与簇模型对应的特征图进行比较,得到最接近的簇模型,也就是数据分布最为相似的簇模型进行推理。

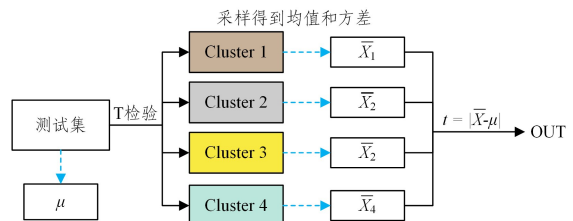


图 4 MCFL 中的客户端

Fig. 4 Client in MCFL

算法 2 MCFL

输入:全局模型 W^* ,簇模型 Cluster,簇的数量 K ,参与训练的客户端数量 N ,客户端模型 w_i

输出:全局模型, K 个簇模型, N 个客户端模型

1. /* MCFL 的流程 */
2. $w_1, w_2, w_3, \dots, w_n \leftarrow W^*$
3. for $t \leftarrow 1$ to T do
 - 3.1. $w_n^i = w_n^{i-1} - \eta \nabla \text{Loss}(w_n^{i-1}, x_n), \forall n \in \{1, 2, 3, \dots, N\}$
 - 3.2. $c_1, c_2, c_3, \dots, c_K \leftarrow K\text{-means}(w_1, w_2, \dots, w_N) / * \Gamma$ 为簇的集合 */
 - 3.3. for $k \leftarrow 1$ to K do
 - 3.3.1. $c_k(x) = \frac{1}{\text{len}(c_k)} \sum_{i=1}^{\text{len}(c_k)} w_i(x)$
 - 3.3.2. $w_1, w_2, w_3, \dots, w_n \leftarrow c_k / *$ 将簇模型发送给每个参与者 */
4. return($W^*, w_1, w_2, \dots, w_N$)

4 实验结果

4.1 实验过程

本文中的实验基于 Mnist 和 EMnist 公共数据集,采用 CNN 网络模型。

Mnist:包含了 250 个不同的人手写的数字构成,一个 10 分类问题,共计 60000 数据,其中 90% 作为训练数据,10% 作为测试数据,采用与 CFL 相同的处理方式^[1],实验中采用 10 个客户端,通过 $\alpha=0.1$ 迪利克雷分布随机采样将数据划分到 10 个客户端,随机选取 5 个客户端的数据翻转,批大小为 30,网络采用两层卷积层、两层池化层、两层 Relu 层和全连接层,最后使用 Softmax 层作为输出。

EMnist:一共包含 80 多万张图片数据集数字,其中 10 类 [0-9]、26 类 [a-z]、26 类 [A-Z],共 62 类,其中 90% 作为训练数据,10% 作为测试数据,采用 20 个客户端,批大小为 64。通过 $\alpha=0.1$ 迪利克雷分布随机采样将数据划分到 20 个客户端,随机选取 10 个客户端的数据旋转 180°。其中模型与 Mnist 采用的模型类似,在中间添加一个 dropout 层以防过拟合。

4.2 实验结果

这部分主要将聚类联邦学习和优化后的结果进行比较,主要针对 CFL 和 IFCA 两种聚类联邦学习进行优化,图 5 给出了 CFL 以及 CFL 的优化方案在 EMnist 数据集上的准确率,在进行 20 轮后开始聚类,可以发现,在聚类之前 3 种方法的簇间学习准确率类似,在聚类后 CFL 簇间学习准确率下降明显,最后准确率达到 0.27 左右,下降了 0.3 左右;采用 AWCFL 方案后,聚类没有对簇间学习效率造成明显的影响,簇间准确率维持在 0.6 左右;在采用 MCFL 化后,簇间准确率达到 CFL 簇内学习效果,准确率提升 0.2,最终达到 0.8 左右的准确率。可以看出在 EMnist 数据集中,MCFL 提升明显,AWCFL 次之,两种方法都能有效地提升 CFL 中的簇间学习率。

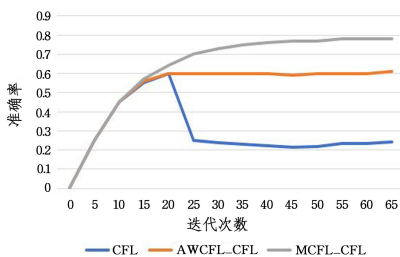


图 5 CFL 及优化方案在 EMnist 上簇间学习率

Fig. 5 Inter cluster learning rate of CFL and optimization on EMnist

图 6 为 CFL 以及优化方案在 EMnist 数据集上的准确率,可以发现在聚类之前 3 种方法的簇间学习准确率类似,在聚类后 CFL 簇间学习准确率下降明显,最后达到 0.4 左右,下降了 0.5 左右;采用 AWCFL 方案后,簇间准确率维持在 0.8 左右;在采用 MCFL 优化后,簇间准确率达到 CFL 簇内学习效果,准确率提升 0.15,最终达到 0.95 左右的准确率。可以得到结论,在 MNIST 数据集中,MCFL 优化提升明显,AWCFL 次之,两种方法都能有效地提升 CFL 中的簇间学习率。

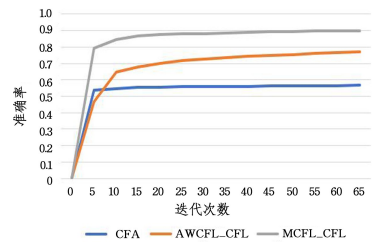


图 6 CFL 及优化方案在 Mnist 上簇间学习率

Fig. 6 Inter cluster learning rate of CFL and optimization on Mnist

表 1 为 AWCFL 中 β 对优化性能的影响,不同的取值会影响簇间和簇内准确率,当取值为 0.5 时,簇间和簇内准确率之和达到最大值。

表 1 β 不同的取值对 AWCFL 的影响

数据集	β 取值	0	0.1	0.25	0.5	1
Mnist	簇间准确率	0.43	0.47	0.69	0.83	0.86
	簇内准确率	0.95	0.92	0.89	0.84	0.43
EMnist	簇间准确率	0.23	0.37	0.49	0.64	0.44
	簇内准确率	0.72	0.70	0.68	0.65	0.23

表 2 为 CFL 以及优化方案多次实验的平均值,可以看出,AWCFL 和 MCFL 优化后相比 CFL 模型簇内准确率下降,簇间准确率大幅度提升,其中 MCFL 准确率最高。

表 2 在 Mnist, EMnist 数据集中 10 次实验平均准确率

Table 2 Average accuracy of 10 experiments in Mnist and EMnist datasets

数据集	簇间/内准确率	FedAvg	CFL	AWCFL	MCFL
Mnist	簇间准确率	0.45	0.45	0.84	0.94
	簇内准确率		0.95	0.83	0.94
EMnist	簇间准确率	0.23	0.24	0.61	0.78
	簇内准确率		0.71	0.61	0.78

图 7 和图 8 分别为 IFCA 以及优化方案在 Mnist 数据集和 EMnist 数据集上的簇间准确率,IFCA 是在训练中持续聚类的过程,它不会像 CFL 中聚类后簇间准确率突然下降,分析发现,在迭代第 5 次后,MCFL_IFCA 优于 IFCA,IFCA 优于 AWCFL_IFCA,AWCFL 方法中训练时簇间聚合会引起训练变慢,训练到后面,MCFL_IFCA 明显优于 IFCA 和 AWCFL_IFCA。

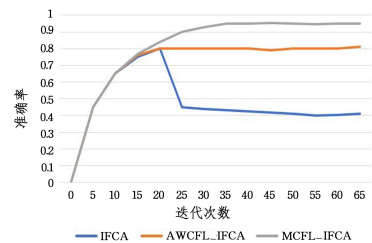


图 7 IFCA 及优化方案在 Mnist 上簇间学习率

Fig. 7 Inter cluster learning rate of IFCA and optimization on Mnist

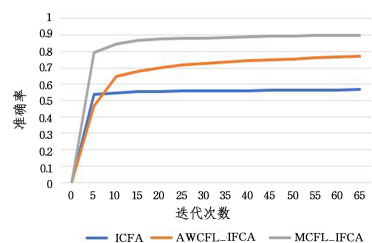


图 8 IFCA 及优化方案在 EMnist 上簇间学习率

Fig. 8 Inter cluster learning rate of IFCA and optimization on EMnist

最后经过实验分析可知,两种方案在簇间准确率提升明显,在 AWCFL 优化中,是以簇内准确率略微下降为代价换取簇间准确率的提升。从表 1 中可以中看出,AWCFL 在 Mnist 中下降 0.12,在 EMnist 中下降 0.06,相比于簇间准确率大幅提升,这样的代价是可以接受的。在 MCFL 优化方案中,为了能够合理选择簇模型,参与者需要上传特征均值,可能存在安全隐患,但是参与者会对数据集样本采样,同时簇内会对参与者上传的特征进行二次平均,通过均值反演出参与者数据集的隐患很小。

结束语 本文致力于解决聚类联邦学习在实践中泛用性低的问题。为此,本文设计了两种优化方案。AWCFL 是在聚类联邦学习中簇内聚合时将其他簇以低权重的方式聚合到簇模型中;MCFL 是将每个簇模型同步到每个客户端,在推理阶段客户端选择合适的簇模型测试。上述的两种方案在 Mnist 和 EMnist 数据集上与 CFL 和 FedAvg 比较准确率,当数据是非独立同分布时,AWCFL 的簇间准确率明显提升但是不及 CFL 的簇内准确率,MCFL 的簇间准确率提升明显,与 CFL 簇内准确率一致。

然而上述优化也付出了一定代价,在 AWCFL 中,服务端需要更多的计算资源用于多次聚合,簇间聚合也会使得簇模型失去学习的独特性,从而导致簇内准确率有一定下降;在 MCFL 中,本地端则需要更多存储模型的空间,对联邦学习的通信要求也较高。

参考文献

[1] MCMAHAN B, MOORE E, RAMAGED, et al. Communication-efficient learning of deep networks from decentralized data[C]// Artificial Intelligence and Statistics. PMLR, 2017:1273-1282.

[2] HARD A, RAO K, MATHEWS R, et al. Federated learning for mobile keyboard prediction[J]. arXiv:1811.03604, 2018.

[3] BRISIMI T S, CHEN R, MELA T, et al. Federated learning of predictive models from federated electronic health records[J]. International Journal of Medical Informatics, 2018, 112:59-67.

[4] SIM K C, ZADRAZIL P, BEAUFAYS F. An investigation into on-device personalization of end-to-end automatic speech recognition models[J]. arXiv:1909.06678, 2019.

[5] MANSOUR Y, MOHRI M, RO J, et al. Three approaches for personalization with applications to federated learning[J]. arXiv:2002.10619, 2020.

[6] MARFOQ O, NEGLIA G, BELLET A, et al. Federated multi-task learning under a mixture of distributions[J]. Advances in Neural Information Processing Systems, 2021, 34:15434-15447.

[7] SMITH V, CHIANG C K, SANJABI M, et al. Federated multi-task learning[J]. arXiv:1705.10467, 2018.

[8] LI T, SAHU A K, TALWALKARA, et al. Federated learning: Challenges, methods, and future directions[J]. IEEE Signal Pro-

cessing Magazine, 2020, 37(3):50-60.

[9] GHOSH A, CHUNG J, YIN D, et al. An efficient framework for clustered federated learning[J]. Advances in Neural Information Processing Systems, 2020, 33:19586-19597.

[10] DUAN M, LIU D, CHEN X, et al. Self-balancing federated learning with global imbalanced data in mobile systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 32(1):59-71.

[11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.

[12] RADOVANOVIC M, NANOPOULOS A, IVANOVIĆ M. On the existence of obstinate results in vector space models[C]// Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2010:186-193.

[13] SATTLER F, MÜLLER K R, SAMEK W. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(8):3710-3722.

[14] SATTLER F, MÜLLER K R, WIEGAND T, et al. On the byzantine robustness of clustered federated learning[C]// IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2020). IEEE, 2020:8861-8865.

[15] WANG L, XU S, WANG X, et al. Addressing class imbalance in federated learning[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2021:10165-10173.

[16] SARKAR S, GHOSH A K. On perfect clustering of high dimension, low sample size data[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42(9):2257-2272.

[17] BRIGGS C, FAN Z, ANDRAS P. Federated learning with hierarchical clustering of local updates to improve training on non-IID data[C]// 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020:1-9.



LI Renjie, born in 1997, postgraduate. His main research interests include federated learning and machine learning.



YAN Qiao, born in 1972, Ph.D, professor, Ph.D supervisor, is a member of China Computer Federation. Her main research interests include network security, cloud computing, software-defined networking, and artificial intelligence.