CaPriDe Learning: Confidential and Private Decentralized Learning Based on Encryption-friendly Distillation Loss

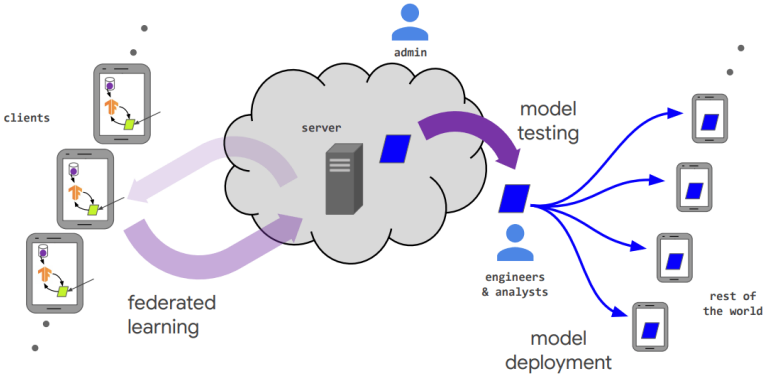
Nurbek Tastan Karthik Nandakumar

MBZUAI, Abu Dhabi, UAE
TUE-PM-379

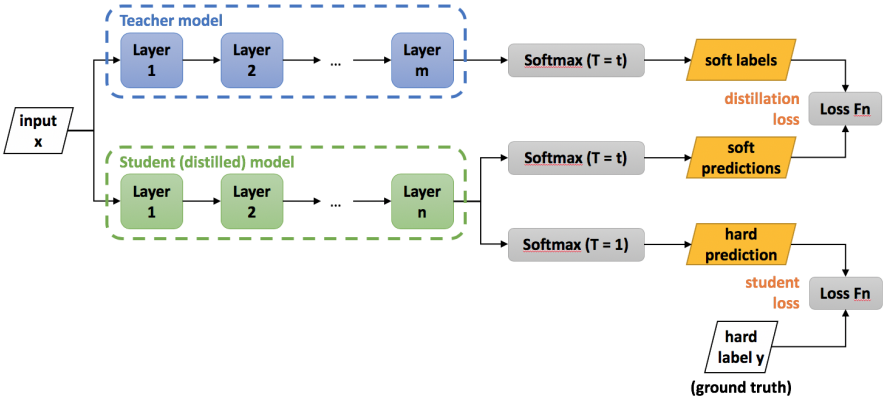
CVPR 2023

Outline

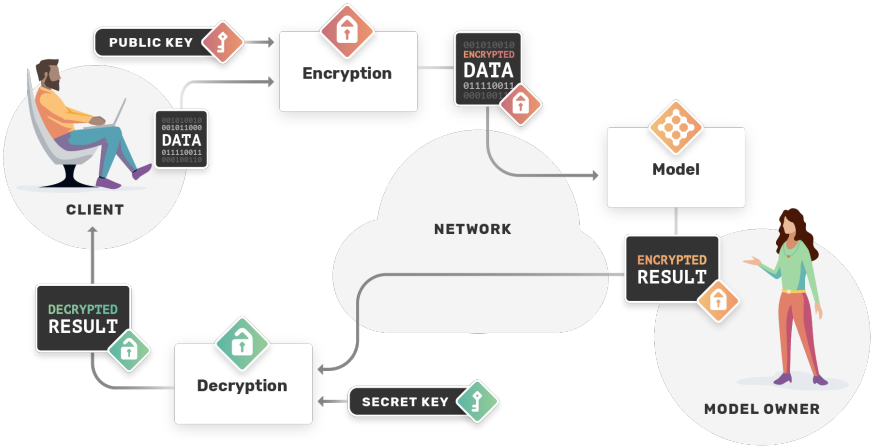
Federated Learning



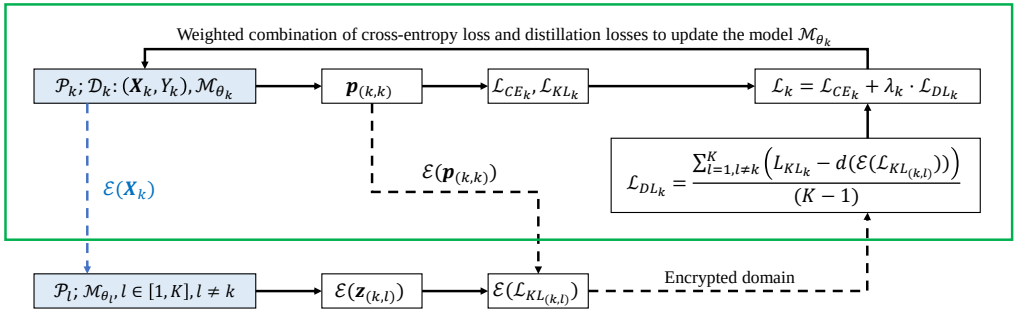
Knowledge Distillation



Homomorphic Encryption



CaPriDe Learning



CaPriDe Learning

The overall loss of $\mathcal{P}_k$

$$\mathcal{L}_k = \mathcal{L}_{CE_k} + \lambda_k \mathcal{L}_{DL_k}$$

The total distillation loss for $\mathcal{P}_k$

$$\mathcal{L}_{DL_k} = \frac{1}{K-1} \sum_{l=1, l \neq k}^{K} \mathcal{L}_{DL_{(k,l)}}$$

Pairwise distillation loss

$$\mathcal{L}_{DL_{(k,l)}} = \sum_{j=1}^{N_k} \left(\boldsymbol{p}_{j,(k,k)} \cdot \log \boldsymbol{p}_{j,(k,k)} \right) - \sum_{j=1}^{N_k} \boldsymbol{p}_{j,(k,k)} \cdot \left(\frac{\boldsymbol{z}_{j,(k,l)}}{T} - \log \left(\sum_{j'=1}^{N_k} \exp \left(\frac{\boldsymbol{z}_{j',(k,l)}}{T} \right) \right) \right)$$

Experimental Setup

Architecture	Dataset	Batch size	λ_k	Description
ResNet-18	CIFAR-10	128	50	60000 (10 classes), 32x32 images
	CIFAR-100	128	50	60000 (100 classes), 32x32 images
	HAM10000	32	20	10015 (7 classes), 224x224 images

Compared methods:

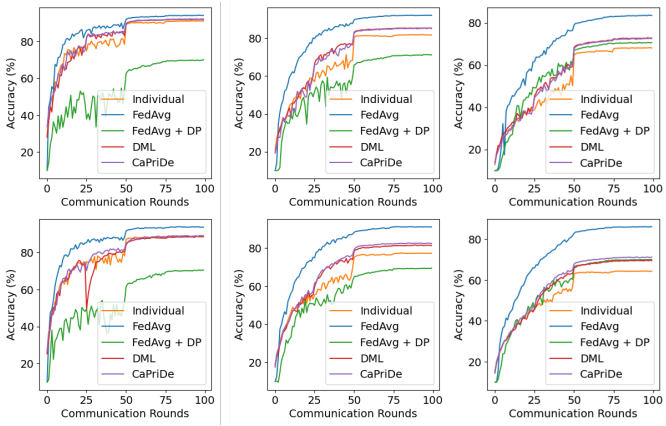
- FedAvg (McMahan *et al.*, 2016)
- FedAvg + DP (Wei *et al.*, 2020)
- DML (Zhang *et al.*, 2017)

Data Partition:

- **Homogeneous:** each participant has an equal number of samples per class;
- **Heterogeneous:** each participant has an unequal number of samples determined randomly (both total number and number of samples per class);
- **Non-overlapping class distribution:** each participant has samples from a non-overlapping subset of classes.

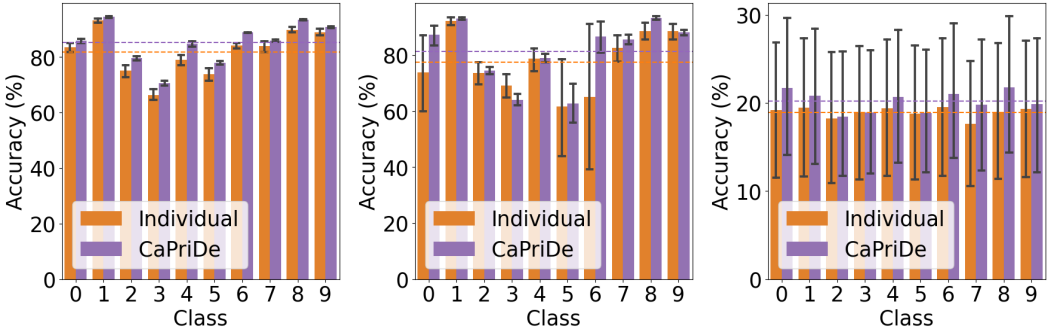
CIFAR-10 Results

Top row:
Homogeneous setting.
Bottom row:
Heterogeneous setting.
Columns: $K = 2, 5, 10$.



Per-class accuracy. $K = 5$.

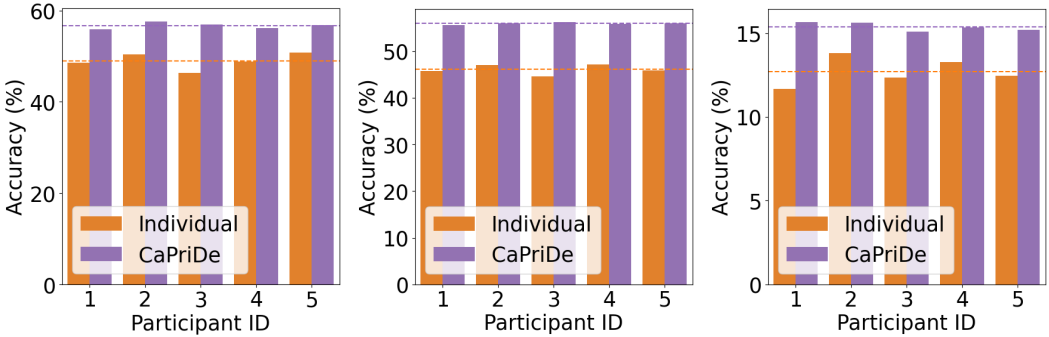
Per-class accuracy of CaPriDe Learning on Cifar-10.



(a) Homogeneous (b) Heterogeneous (c) No overlap

Per-participant accuracy. $K = 5$.

Per-participant accuracy of CaPriDe Learning on Cifar-100.



(a) Homogeneous (b) Heterogeneous (c) No overlap

Comparison with L_2 loss

Collaboration gain of CaPriDe learning algorithm with the proposed approximate KL loss in comparison with the L_2 loss. As a distillation loss, the proposed approximate KL divergence loss outperforms the L_2 loss by far.

Setting	K	Individual	CaPriDe (KL)	CaPriDe (L_2)
Homogeneous	2	91.050	**92.155**	91.025
	5	81.770	**85.194**	82.710
	10	68.065	**72.580**	68.272
Heterogeneous	2	87.485	**88.424**	87.320
	5	77.336	**81.324**	76.070
	10	64.320	**70.520**	65.010

FHE Results

FHE configuration parameters and memory / computational requirements for the chosen datasets based on ResNet-18 architecture.

	CIFAR-10	CIFAR-100	HAM10000
Security Level	128	128	128
Number of slots	16384	16384	16384
Time taken to encrypt one sample	90 ms	103 ms	619 ms
Ciphertext size of one sample	29.101 KB	29.152 KB	1.359 MB
Time taken to encrypt a batch of 32 samples	1.31 s	1.26 s	15.57 s
Encrypted inference of a batch of 32 samples	110.21 s	112.09 s	896.12 s

Conclusion

- A new collaborative learning algorithm, which exploits encrypted inference and knowledge distillation to achieve confidentiality and privacy without any central orchestration and any need for non-private shared data.
- An encryption-friendly distillation loss that estimates the approximate KL divergence between model predictions and design a protocol to securely compute the loss in the encrypted domain.
- Extensive experiments on the benchmark datasets demonstrate that CaPriDe learning achieves better privacy-utility trade-off among the compared methods.

Code: `https://github.com/tnurbek/capride-learning`