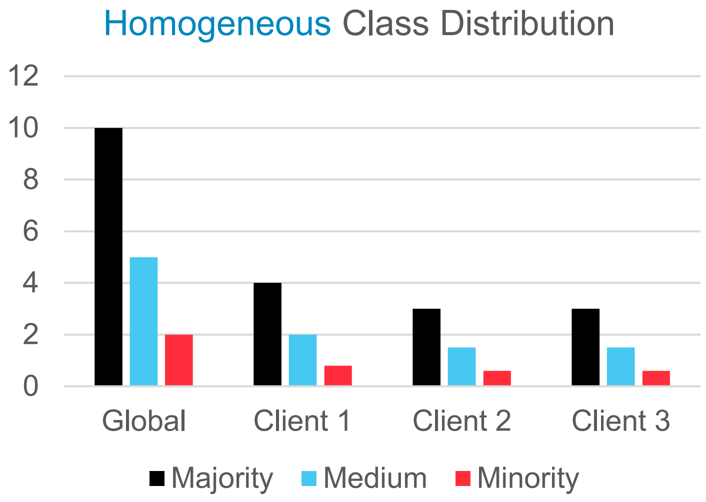
FedReLa: Imbalanced Federated Learning via Re-Labeling

Guangzheng Hu, Patricia Menendez, Feng Liu, Mingming Gong, Guanghui Wang, Liuhua Peng

University of Melbourne

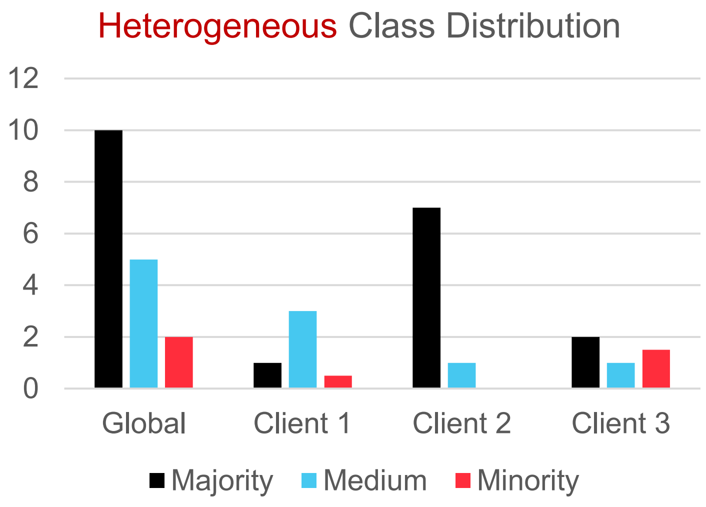
Imbalanced Federated Learning

Heterogeneous + Imbalanced Data Distribution



Rise New Challenges:

- No global class priors
- No local data / class priors exchanges
- Class absence
- Local computation is limited
- Communicational cost
- Global-local distribution mismatch



No data-level method in Federated Learning:

- **Requires global class priors knowledge**: need to identify minority first
- Some clients has **zero seed samples**
- Fixed hyperparameters can not deal with heterogeneous data
- Misaligned multiple biased decision boundaries

Correcting Aggregated Decision Boundary

Parameter Aggregation Perspective: FedAvg, FedNova, FedProx, MOOM, CLIMB

Decision Boundary Aggregation Perspective:
- Local model converges on local data
- Changing local data distribution directly rectifies local decision boundaries

For a local dataset $\widetilde{\mathcal{D}}^{(k)}$, and classes j and ℓ , we define

- Local sample relabelling probabilities:

$$\rho^{(k)}_{\ell\to j}(x) = \Pr(\widetilde{Y}^{(k)} = j \mid X^{(k)} = x, Y^{(k)} = \ell)$$

$$\rho^{(k)}_{j\to \ell}(x) = \Pr(\widetilde{Y}^{(k)} = \ell \mid X^{(k)} = x, Y^{(k)} = j)$$

- Local decision boundaries with relabelling

$$\widetilde{S}^{(k)} = \left\{ x^* \in \mathcal{X} : \frac{P_j(x^*)}{P_\ell(x^*)} = \frac{1-2\rho^{(k)}_{\ell\to j}(x^*)}{1-2\rho^{(k)}_{j\to\ell}(x^*)} \cdot \frac{\pi^{(k)}_\ell}{\pi^{(k)}_j} \right\}$$

- Global Aggregated decision boundary on relabelled local datasets

$$\widetilde{S}^{[w]} = \left\{ x^* \in \mathcal{X} : \frac{P_j(x^*)}{P_\ell(x^*)} = \frac{\sum_{k=1}^K w_k \pi^{(k)}_\ell [1 - 2\rho^{(k)}_{\ell\to j}(x)]/\pi_\ell}{\sum_{k=1}^K w_k \pi^{(k)}_j/\pi_j} \cdot \frac{\pi_\ell}{\pi_j} \right\}$$

- With

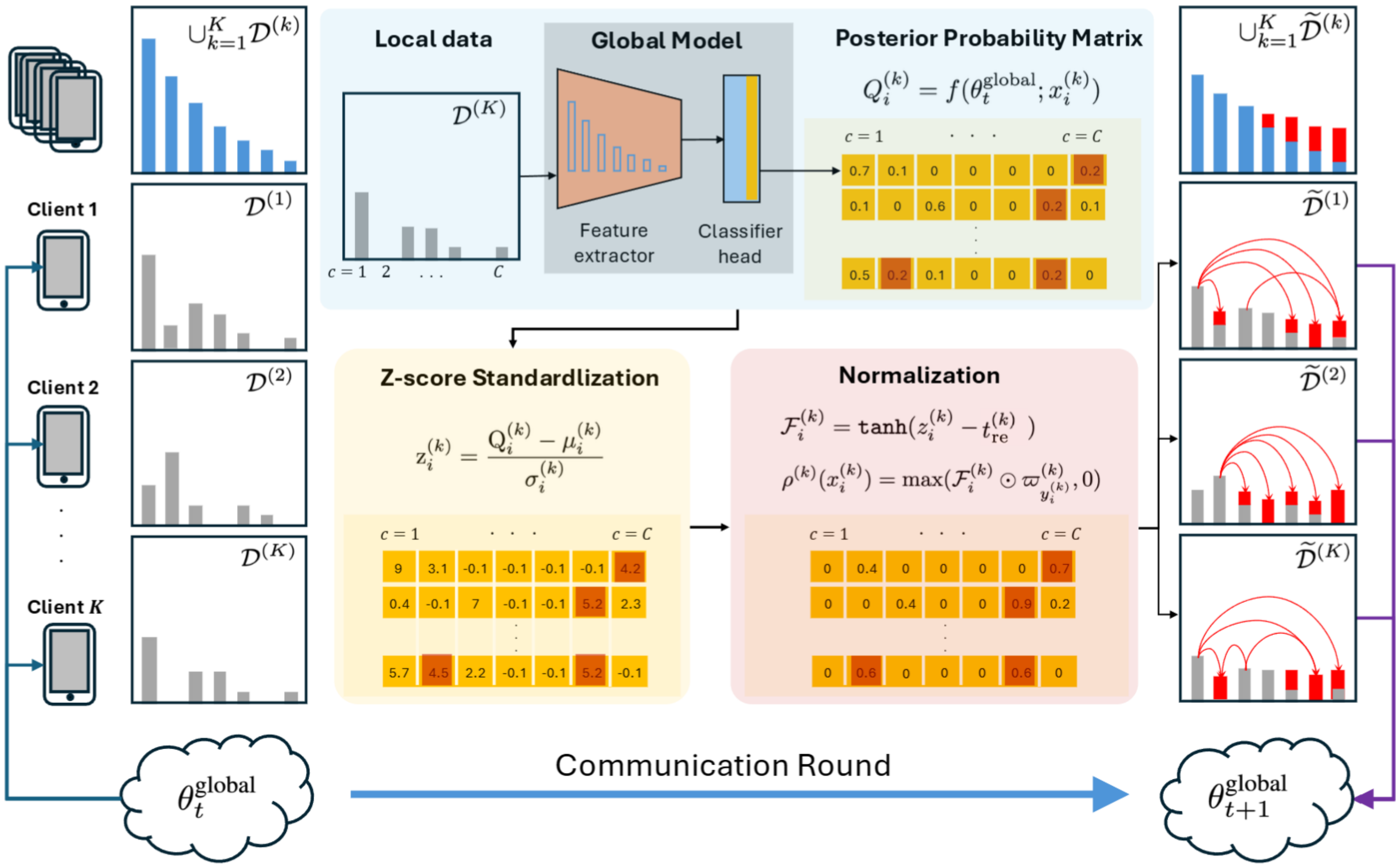
$$w_k = |\mathcal{D}^{(k)}|/|\mathcal{D}|, \sum_{k=1}^K w_k \pi^{(k)}_j/\pi_j = 1.$$

- We have proven that

$$\sum_{k=1}^K w_k \pi^{(k)}_\ell [1 - 2\rho^{(k)}_{\ell\to j}(x)]/\pi_\ell < 1$$

when $\rho^{(k)}_{\ell\to j}(x) > 0$ for all $k \in \{1, \dots, K\}$

FedReLa: Adaptive Re-Labeling



Adaptive Threshold:

$$t_{re,j}^{(k)} = Quantile_{1-\tau}\left(\mathbf{Z}_{:,j}^{(k)}\right), \quad j = 1, \ldots, C,$$

- Localized relabelling
- Safeguarded relabelling

Adaptive Direction of Relabelling:

$$\varpi^{(k)} = 1 - \mathrm{minmax}(n_{\mathcal{Y}}^{(k)})$$

$$\varpi_j^{(k)} = \max(\varpi^{(k)} - \varpi^{(k)}[j], 0)$$

- Only majority to minority
- Localized balancing

FedReLa: Adaptive Re-Labeling

Algorithm 1 `Re-Allocator` re-labels local data

Input: local datasets $\mathcal{D}^{(k)} = \{(x_i^{(k)}, y_i^{(k)})\}_{i=1}^{n_k}$, classifier $f(\cdot;\cdot)$, global model θ_t^{global}

Parameters: Threshold $t_{\text{re}}^{(k)}$, re-labeling round T_{relabel}

for client $k \in K$ at communication round $t = T_{\text{relabel}}$ **do**:

  Compute $\varpi^{(k)}$ by (3) with $n_{\mathcal{Y}}^{(k)}$

  for $(x_i^{(k)}, y_i^{(k)}) \in \widetilde{\mathcal{D}}^{(k)}$ **do**:

    $\mathrm{Q}_i^{(k)} \leftarrow f(\theta_t^{\text{global}}; x_i^{(k)})$

  for $(x_i^{(k)}, y_i^{(k)}) \in \mathcal{D}^{(k)}$ **do**:

    Compute $z_i^{(k)}$ by (3) with $\mathrm{Q}_i^{(k)}$

    $\varpi_{y_i^{(k)}}^{(k)} \leftarrow \mathtt{max}(\varpi^{(k)} - \varpi^{(k)}[y_i^{(k)}], 0)$

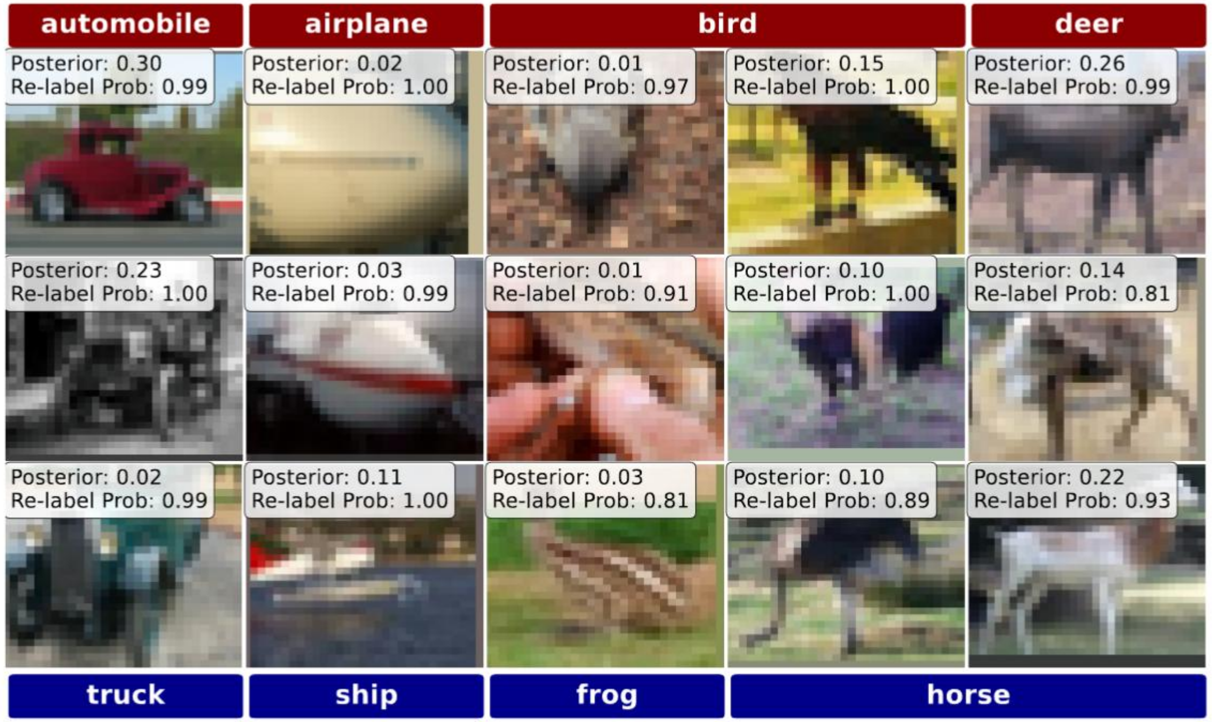
    Compute $\rho^{(k)}(x_i^{(k)})$ by (4)

    $\mathcal{U} \in \mathbb{R}^{|\mathcal{Y}|} \leftarrow \mathtt{Bernoulli}(\rho^{(k)}(x_i^{(k)}))$

    if $\mathcal{U}$ contains 1 **then**:

      $\widetilde{y}_i^{(k)} \leftarrow \mathcal{Y}[\mathtt{argmax}(\rho^{(k)}(x_i^{(k)}))]$

  return $\widetilde{\mathcal{D}}^{(k)} \leftarrow \{(x_i^{(k)}, \widetilde{y}_i^{(k)})\}_{i=1}^{n_k}$



THE UNIVERSITY OF MELBOURNE

Results

THE UNIVERSITY OF MELBOURNE

Dataset	IF	Heterogeneity Method/Metrics	$\alpha = 0.1$ H/M/T-shots	$\alpha = 0.1$ Overall	$\alpha = 0.3$ H/M/T-shots	$\alpha = 0.3$ Overall	$\alpha = 10$ H/M/T-shots	$\alpha = 10$ Overall
CIFAR-10	50	FedETF	85.07/41.03/5.40	47.96	79.60/67.93/43.10	65.15	88.77/73.22/56.81	74.52
		+(FedReLa)	78.70/65.20/35.57	**61.71**	74.95/71.57/55.63	**68.14**	86.80/76.62/71.59	**79.18**
		FedLOGE	56.51/58.98/59.74	58.22	82.81/76.90/67.98	76.59	88.32/81.32/72.40	81.45
		+(FedReLa)	69.55/70.13/53.09	**64.78**	80.09/78.47/76.86	**78.63**	83.79/83.32/84.12	**83.75**
	100	FedETF	66.90/36.80/24.40	40.87	72.05/37.13/25.32	42.88	92.16/69.52/50.13	68.56
		+(FedReLa)	57.60/44.53/36.95	**45.42**	72.59/40.11/30.47	**46.00**	90.17/70.73/66.91	**75.04**
		FedLOGE	44.93/50.27/41.30	45.08	86.71/67.52/57.89	69.43	91.90/75.07/62.50	75.09
		+(FedReLa)	70.60/59.60/57.75	**62.16**	78.62/68.10/69.39	**71.77**	84.58/75.48/79.70	**79.90**
CIFAR-100	50	FedETF	62.97/44.65/20.12	42.35	68.89/46.87/20.28	44.83	71.11/48.78/18.80	44.83
		+(FedReLa)	56.81/48.60/29.20	**44.71**	61.79/51.20/30.38	**47.41**	56.37/53.09/34.30	**47.18**
		FedLOGE	32.92/37.70/34.64	35.09	59.40/47.76/31.40	45.87	68.37/52.31/26.45	47.88
		+(FedReLa)	51.22/46.82/30.37	**42.73**	63.93/51.99/30.04	**48.25**	67.65/51.30/29.77	**48.57**
	100	FedETF	64.48/42.64/13.72	37.11	68.75/48.13/14.93	39.63	71.76/47.48/16.45	39.67
		+(FedReLa)	56.10/46.73/24.14	**40.19**	61.01/51.24/25.02	**42.70**	56.05/50.78/29.86	**42.51**
		FedLOGE	27.00/36.41/25.62	29.21	50.80/43.67/25.26	37.75	69.55/50.40/22.98	42.87
		+(FedReLa)	49.25/47.98/23.11	**38.08**	62.36/52.15/23.98	**42.89**	66.28/50.43/25.92	**43.36**

Table 2. Test accuracies (in %) of different methods on long-tailed CIFAR-10/100.

Method	**50 Clients**		**100 Clients**	
	Overall Accuracy (%)	**H/M/F Accuracy (%)**	**Overall Accuracy (%)**	**H/M/F Accuracy (%)**
FedLC (Zhang et al., 2022)	32.81	56.20/32.42/9.74	23.72	46.41/20.74/4.23
+FedReLa	**34.76**	49.92/**35.24/17.21**	**25.13**	39.82/**28.34/7.02**
FedYoYo (Yan et al., 2025)	40.89	54.32/41.95/24.12	30.73	34.12/29.64/27.92
+FedReLa	**41.31**	53.62/**42.24/25.91**	**32.13**	33.90/**32.72/29.22**
FedETF	31.89	60.67/40.45/9.22	28.71	59.00/33.72/9.73
+FedReLa	**33.94**	55.12/**43.61/15.12**	**31.50**	52.21/**38.73/16.53**
FedLoGe	34.83	57.12/42.21/17.29	33.08	62.71/38.78/13.75
+FedReLa	**35.62**	57.01/**44.32/18.00**	**34.32**	59.90/**42.22/16.00**

Method	**IF**	**Med+Tail Gain (%)**			**Overall Gain (%)**		
+FedReLa		α=0.1	α=0.3	α=10	α=0.1	α=0.3	α=10
FedETF	50	**+54.34**	+16.17	+18.18	**+13.75**	+2.99	+4.66
FedETF	100	**+20.28**	+8.13	+17.99	+4.55	+3.12	**+6.48**
FedLoGe	50	+4.50	+10.45	**+13.72**	**+6.56**	+2.04	+2.30
FedLoGe	100	**+25.78**	+12.08	+17.61	**+17.08**	+2.34	+4.81