

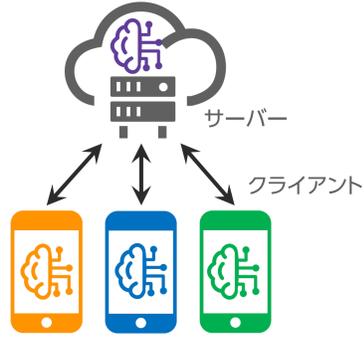
# 連合学習の通信量を削減するための Energyスコアを利用した知識蒸留手法の検討

東 桔也 宮崎 智 大町真一郎 (東北大)

## 背景・目的

### 連合学習 (Federated Learning)

- 分散型の機械学習フレームワーク
- 学習データではなくモデルのパラメータを集約してグローバルなモデルを学習



✓ プライバシー保護

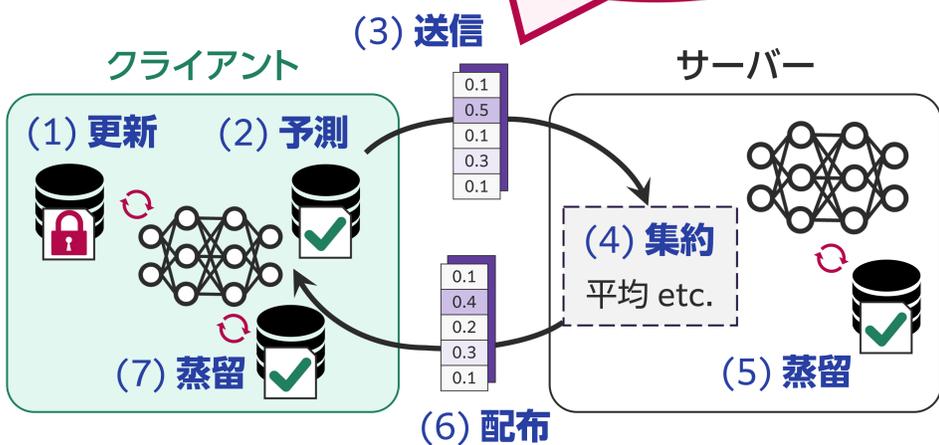
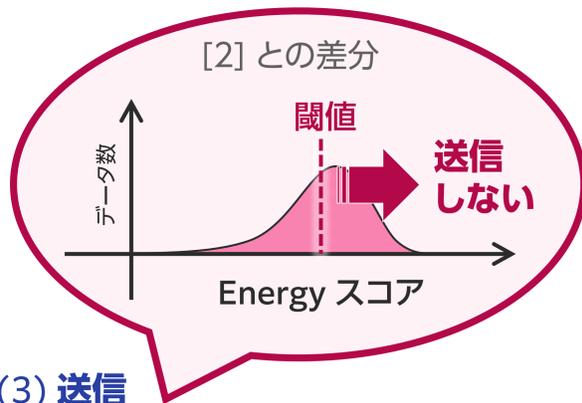
知識蒸留<sup>[1]</sup>を利用してモデル出力のみを集約し、通信コストを大幅に削減する手法<sup>[2]</sup>もある

✗ クライアントが所持するデータが偏っている場合 (Non-IID) に学習の収束が遅い

💡 各クライアントで十分に学習されたデータに対する予測のみを送りたい

## 提案手法

送信する予測に  
閾値を設定する



🔒 : Private (ラベルあり)    ✓ : Public (ラベルなし)

## Energyスコア

$$E(\mathbf{x}; f) = -\log \sum_i^K e^{f_i(\mathbf{x})}$$

$\mathbf{x}$ : 入力     $f$ : 分類器     $K$ : 出力の次元数

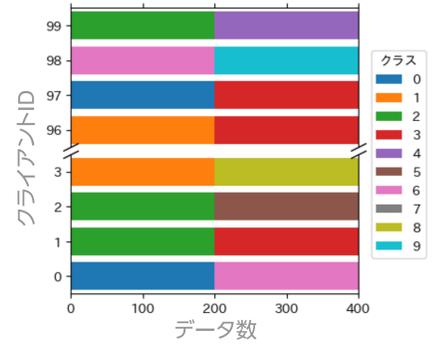
↑ 分布外検出でSoftmax信頼度スコアを上回る精度<sup>[3]</sup>

データサイズを削減しながら、より正確な予測をサーバーに送信して学習を効率化

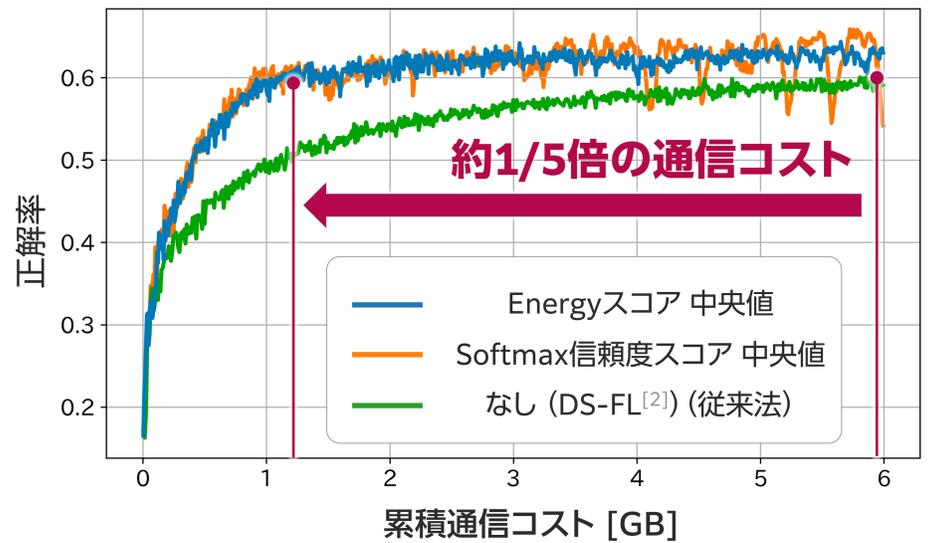
## 実験結果・考察

データセット: CIFAR-10  
クラス数: 10  
クライアント数: 100

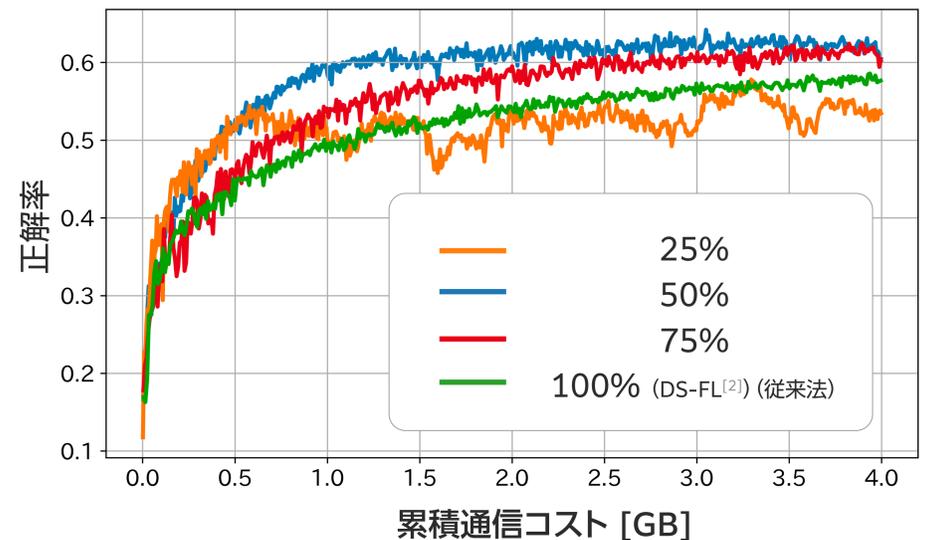
偏りが強くなるように分配



### 実験1 閾値の種類



### 実験2 閾値 (Energyスコア) の大小



閾値大 → 不正確な予測を多く含む  
閾値小 → 蒸留に有益な予測が送信されない

## 結論

偏りが強いNon-IIDデータの場合に、Energyスコアの閾値を利用することで連合学習を効率化し、通信コストを大きく削減できた

## 参考文献

- [1] Hinton, G. et al. Distilling the Knowledge in a Neural Network. *arXiv preprint arXiv:1503.02531*, 2015.
- [2] Itahara, S. et al. Distillation-Based Semi-Supervised Federated Learning for Communication-Efficient Collaborative Training With Non-IID Private Data. In *IEEE Transactions on Mobile Computing*, 2021.
- [3] Liu, W. et al. Energy-based Out-of-distribution Detection. *arXiv preprint arXiv: 2010.03759*, 2021.

プロジェクトページでソースコードと  
研究の概要・詳細を公開しています！

✉ azuma.kitsuya.s5@dc.tohoku.ac.jp

