FedMD: Heterogenous Federated Learning via Model Distillation

NeurIPS 2019
Daliang Li etc.
Cite: 50

2021.03.18.목
발표자: 임진혁

Index

임진혁<eeplearning@units.ac.kr>

FL 논문들의 실효성에 대한 의문

- FedAvg 논문의 낮은 contribution
- Federate Learning의 실체가 궁금하다

임진혁<eeplearning@units.ac.kr>

Introduction: 논문 선택 이유

FL's Goal: 데이터가 분산된 상태로 학습한 성능 ~= 한 곳에 집적된 상태로 학습한 성능

Benefit

- **Data Collecting Cost 회피 가능**
- **Data Privacy(보안)**

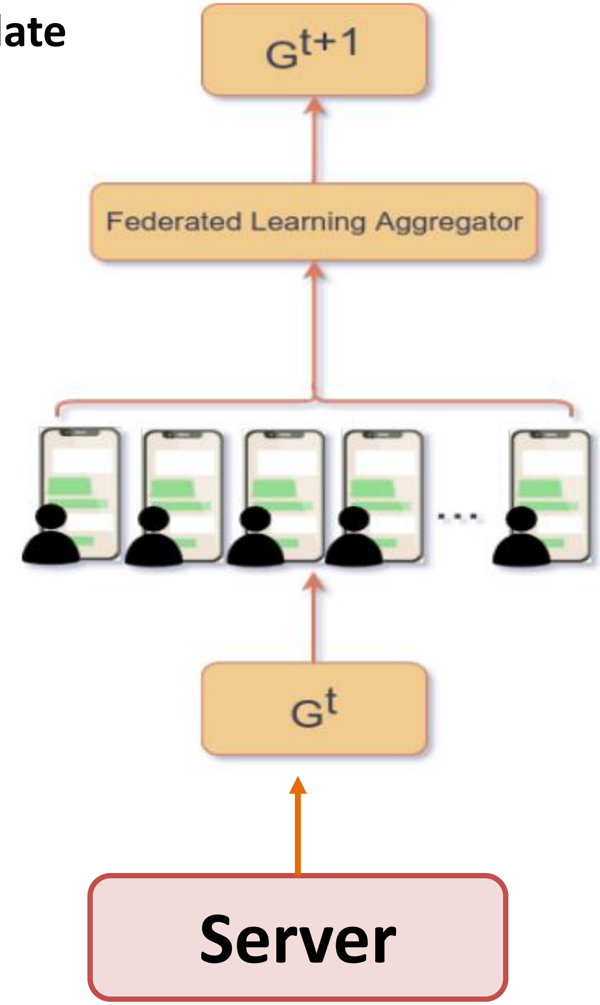
임진혁<eeplearning@units.ac.kr>

2 MIIL Related: Federate Learning

FedAVG : Client들이 SGD를 통해 학습하고 업데이트 정보(LOSS) Aggregate => Global update

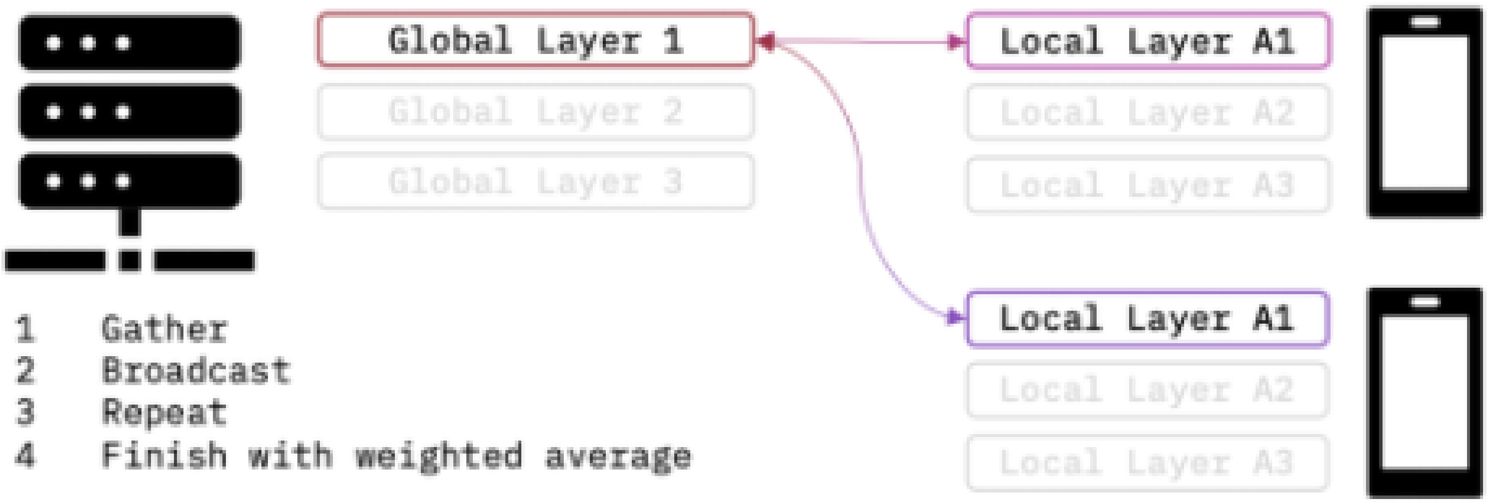
Algorithm 2 **Generalized FedAvg**

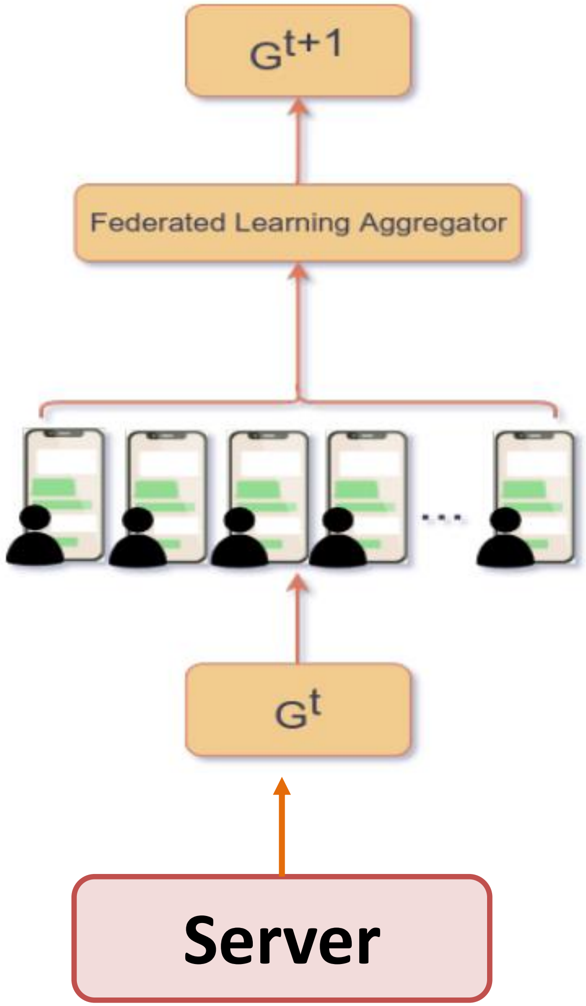
Initialization: x_0
for $t = 0, \cdots, T-1$ **do**
  Sample subset $\mathcal{S}$ of clients
  $x_{i,0}^t = x_t$
  for each client $i \in \mathcal{S}$ **in parallel do**
    for $k = 0, \cdots, K-1$ **do**
      Compute an unbiased estimate $g_{i,k}^t$ of $\nabla F_i(x_{i,k}^t)$
      $x_{i,k+1}^t = \textsc{ClientOpt}(x_{i,k}^t, g_{i,k}^t, \eta_l, t)$
    $\Delta_i^t = x_{i,K}^t - x_t$
  $\Delta_t = \frac{1}{|\mathcal{S}|}\sum_{i\in\mathcal{S}} \Delta_i^t$
  $x_{t+1} = \textsc{ServerOpt}(x_t, -\Delta_t, \eta, t)$



Related: Federate Learning

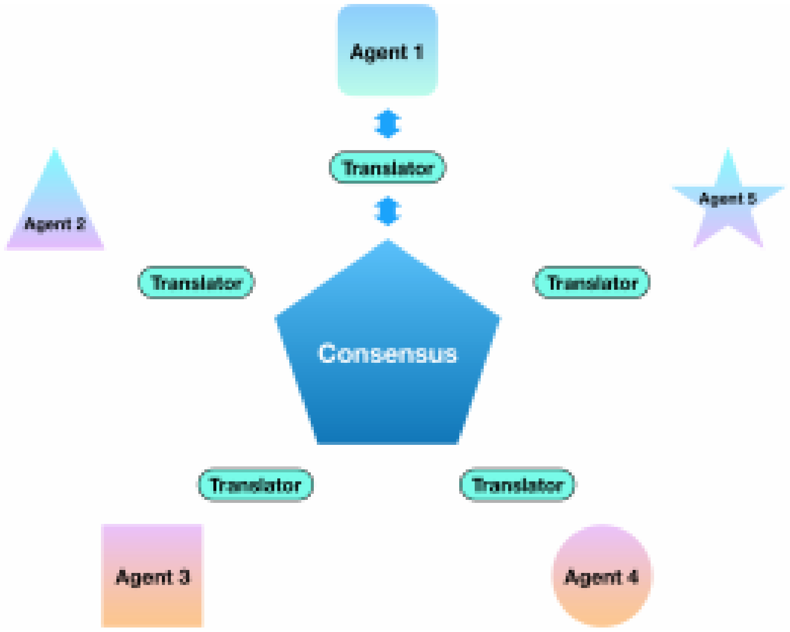
FedMA : 모델의 Layer을 공유 (layer-wise matching)

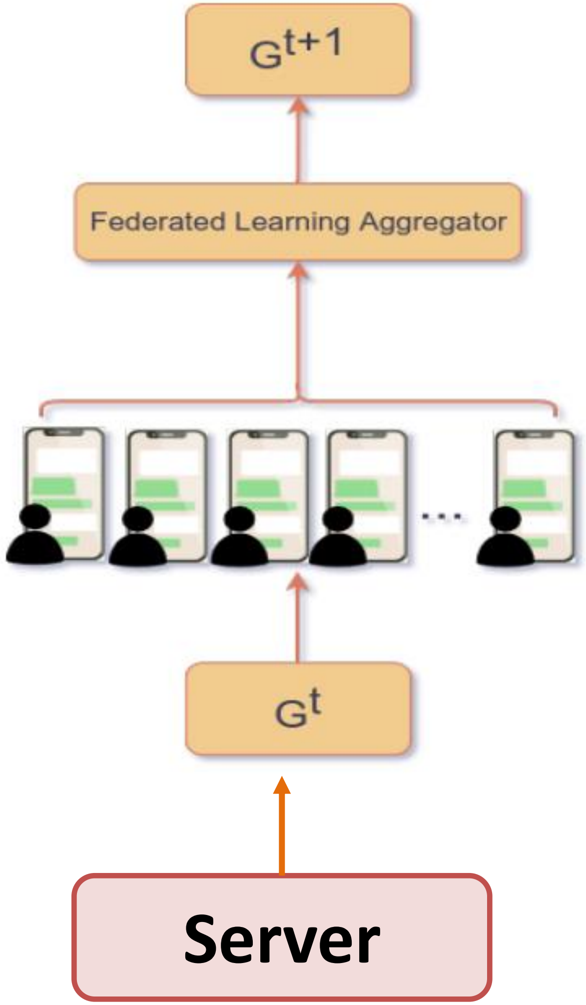




2 Related: Federate Learning

FedMD : KD를 활용하여 Federate





임진혁<eeplearning@unist.ac.kr>

Both of FL's benefits & challenges come from same setting

임진혁<eeplearning@units.ac.kr>

Introduction: 논문 선택 이유

Both of FL's benefits & challenges come from same setting

⇒ Heterogeneity(Data distributed on different clients)

임진혁<eeplearning@units.ac.kr>

Both of FL's benefits & challenges come from same setting

⇒ Heterogeneity(Data distributed on different clients)

- Cloud storage cost efficiency
- Privacy
- Client Drifts
  - Bandwidth heterogeneity
  - Computational power heterogeneity
  - Statistical heterogeneity (the non i.i.d problem)

임진혁<eeplearning@units.ac.kr>

Introduction: 논문 선택 이유

FL's Goal: 데이터가 **분산된 상태로** 학습한 성능 ~= 한 곳에 집적된 상태로 학습한 성능

임진혁<eeplearning@units.ac.kr>

Introduction: 논문 선택 이유

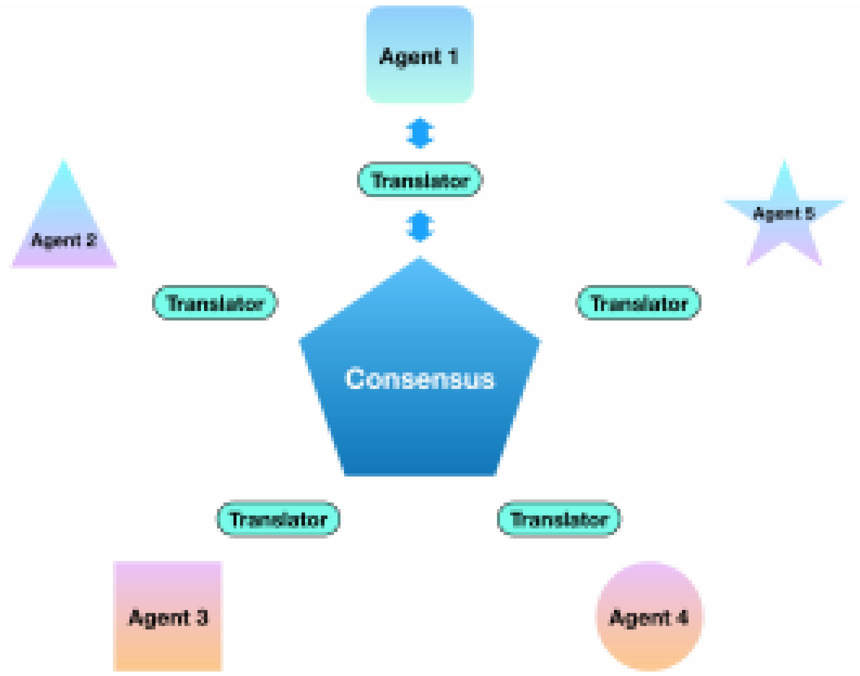
FL's Goal: heterogenity에 기반하는 challenges 들을 극복하면서 Centralized Model의 성능을 구현

- Communication cost efficiency
- Heterogeneity / Non- iid 문제를 해결할 수 있어야 한다.
- Privacy가 보장되어야 한다.
- Single Storage Model의 성능과 비슷해야 한다.
- Client(edge)가 운용할 수 있는 크기 어야 한다.

임진혁<eeplearning@units.ac.kr>

FedMD's Contribution

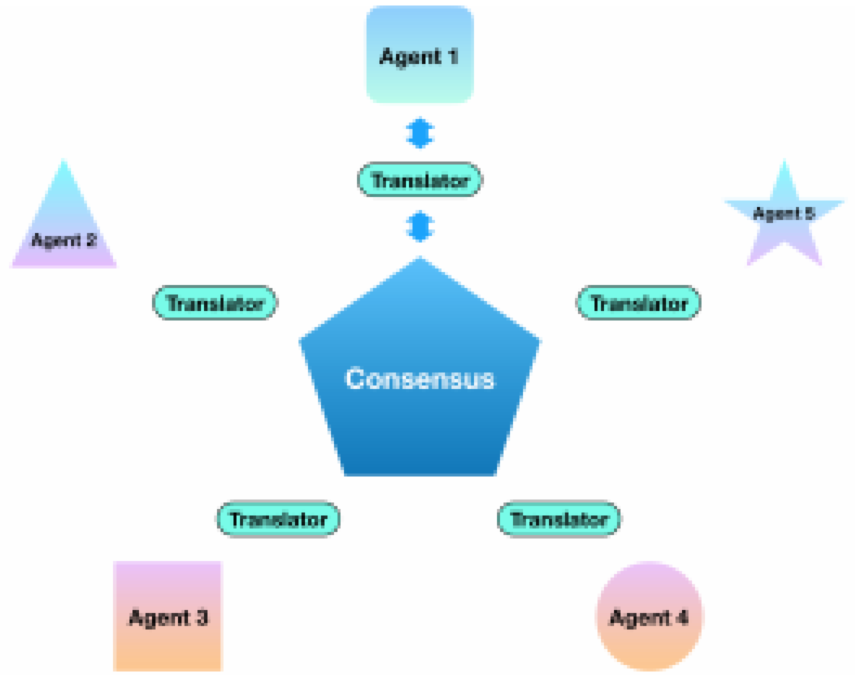
=> **Model Heterogeneity**



FedMD’s Contribution

=> Model Heterogeneity

- Transfer Learning
- Knowledge Distillation



임진혁<eeplearning@units.ac.kr>

Proposed Method: Why We need Model Heterogeneity?

FedAvg

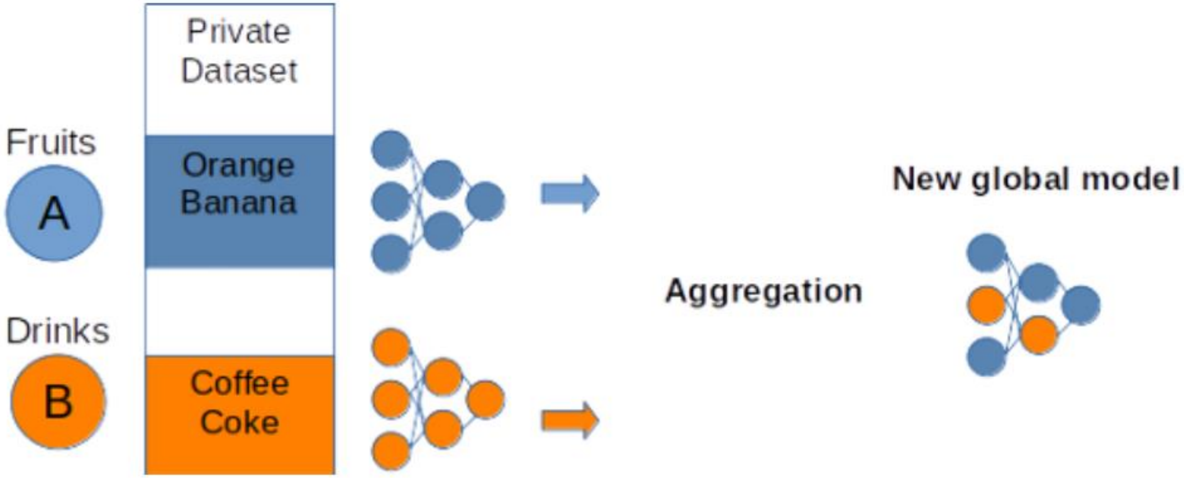


Fig 1. FedAvg: Simple model averaging (Image by Author)

Proposed Method: Why We need Model Heterogeneity?

FedAvg

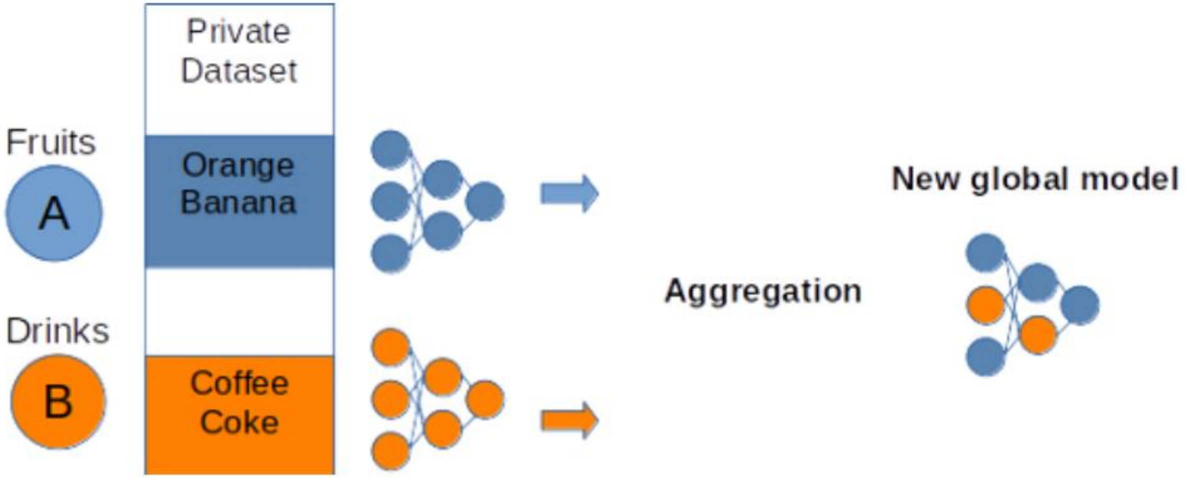
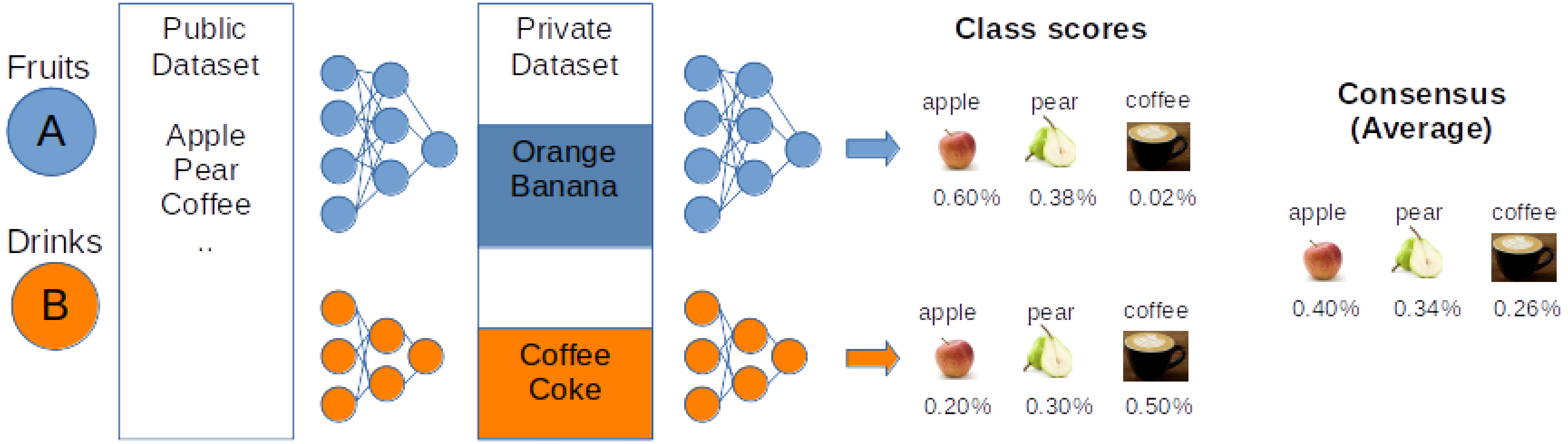


Fig 1. FedAvg: Simple model averaging (Image by Author)

FedMd



임진혁<eeplearning@unist.ac.kr>

We need **Model Heterogeneity**

Why?

임진혁<eeplearning@unist.ac.kr>

We need Model Heterogeneity

Why?

Actually..

it’s natural that using fixed architecture of a centralized model

but..

임진혁<eeplearning@unist.ac.kr>

We need Model Heterogeneity

- Realistic Business Settings
- Client Heterogeneity
- Specification

  (Model for rich information)

  (Different task , different model)

How to Model & Data Hetero. at same time?

모델이 얻은 지식을 모델의 파라미터 말고 데이터를 전달하는 방식이 무엇일까?

⇒ KD

⇒ KD라면 필요한 것은?

⇒ Base Dataset

임진혁<eeplearning@unist.ac.kr>

Proposed Method: FedMD

Server

Base Dataset

Translator: 지식의 통로역할을 할 public data

Unique Model

각 client의 독자적인 모델 ⇔ *Global Model (Federate Setting)*

Private Dataset

Unique Model이 fine-tune 하는 데이터

임진혁<eeplearning@unist.ac.kr>

(1) Transfer Learning

(2) Collaboration Phase

임진혁<eeplearning@unist.ac.kr>

Proposed Method: FedMD

(1) Transfer Learning

- Train base dataset
- Train private dataset(transfer learning)

(2) Collaboration Phase

- Communicate (KD)
- Aggregate
- Distribute
- Digest
- Revisit

임진혁<eeplearning@unist.ac.kr>

(2) Collaboration Phase

Communicate(KD)

=> **Class Score** (KD의 Soft label) on the base dataset

임진혁<eeplearning@unist.ac.kr>

(2) Collaboration Phase

Aggregate

=> Mean[Class score] (updated consensus)

Distribute (Download)

Digest

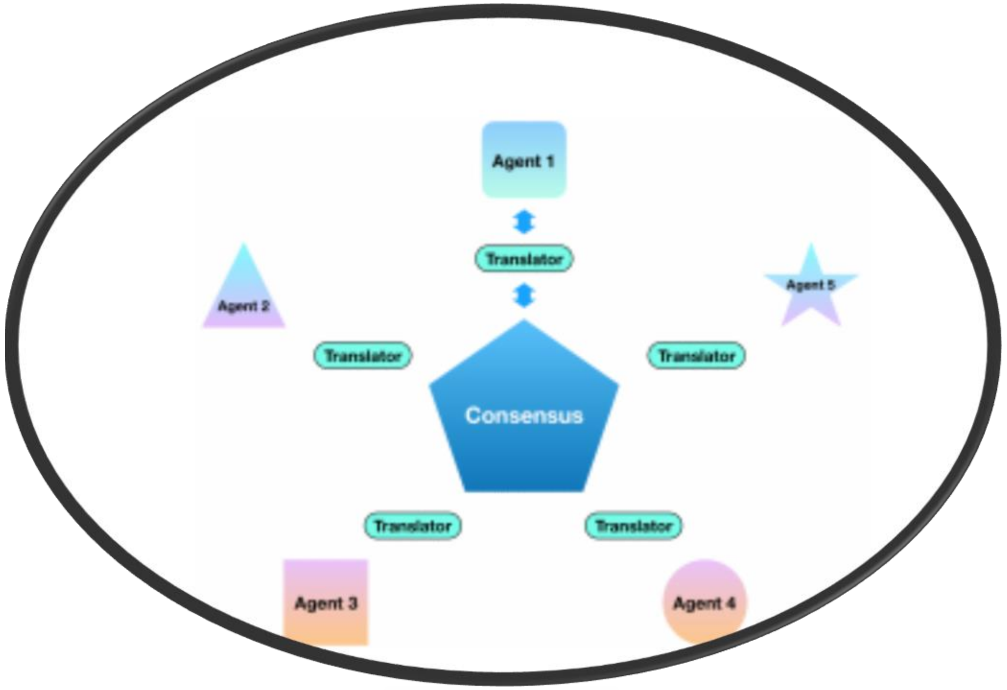
=> Train unique model to label: (updated consensus)

임진혁<eeplearning@unist.ac.kr>

(2) Collaboration Phase

Algorithm 1: The FedMD framework enabling federated learning for heterogeneous models.

Input: Public dataset $\mathcal{D}_0$, private datasets $\mathcal{D}_k$, independently designed model f_k, $k = 1 \ldots m$,
Output: Trained model f_k
Transfer learning: Each party trains f_k to convergence on the public $\mathcal{D}_0$ and then on its private $\mathcal{D}_k$.
for j=1,2...P **do**
  Communicate: Each party computes the class scores $f_k(x_i^0)$ on the public dataset, and transmits the result to a central server.
  Aggregate: The server computes an updated consensus, which is an average $\tilde{f}(x_i^0) = \frac{1}{m}\sum_k f_k(x_i^0)$.
  Distribute: Each party downloads the updated consensus $\tilde{f}(x_i^0)$.
  Digest: Each party trains its model f_k to approach the consensus $\tilde{f}$ on the public dataset $\mathcal{D}_0$.
  Revisit: Each party trains its model f_k on its own private data for a few epochs.
end



임진혁<eeplearning@unist.ac.kr>

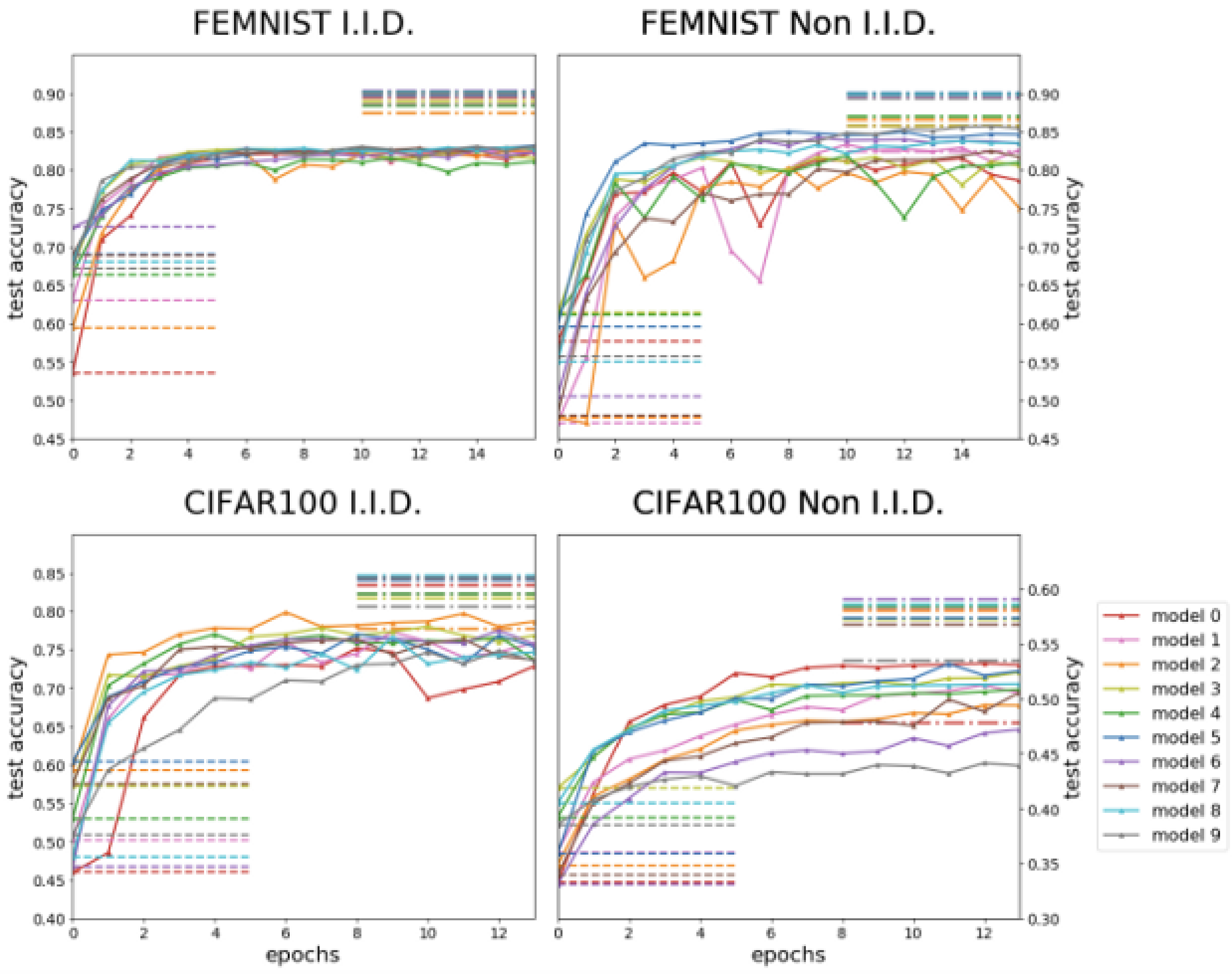
Experiment on 10 participants
Use subset for practical purpose

(1)MNIST / FEMNIST

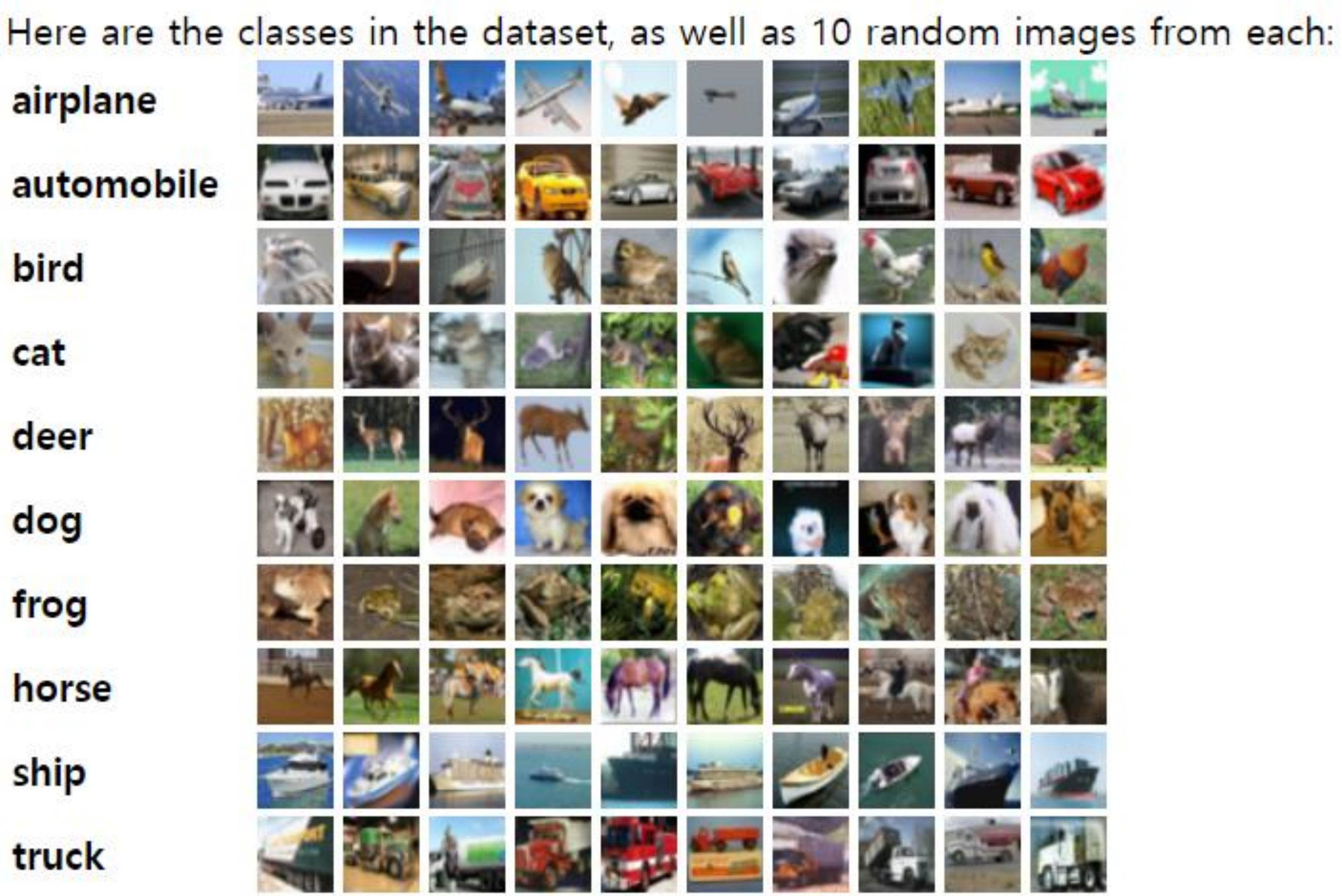
(2)CIFAR10 / CIFAR100

Model Heterogeneity

CNN With Different layer / filter



CIFAR10

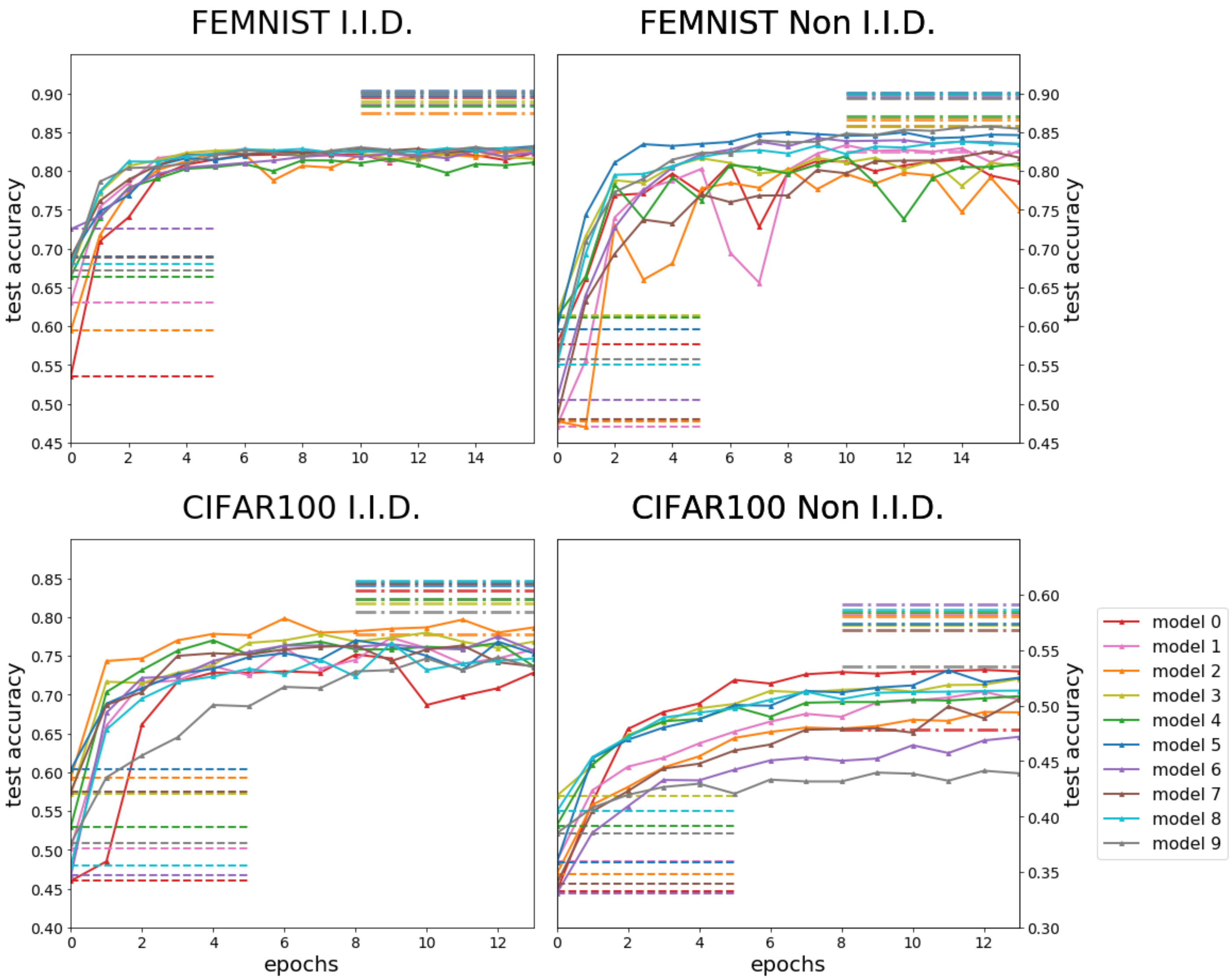

Here are the classes in the dataset, as well as 10 random images from each:
airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

CIFAR100

Superclass	Classes
aquatic mammals	beaver, dolphin, otter, seal, whale
fish	aquarium fish, flatfish, ray, shark, trout
flowers	orchids, poppies, roses, sunflowers, tulips
food containers	bottles, bowls, cans, cups, plates
fruit and vegetables	apples, mushrooms, oranges, pears, sweet peppers
household electrical devices	clock, computer keyboard, lamp, telephone, television
household furniture	bed, chair, couch, table, wardrobe
insects	bee, beetle, butterfly, caterpillar, cockroach
large carnivores	bear, leopard, lion, tiger, wolf
large man-made outdoor things	bridge, castle, house, road, skyscraper
large natural outdoor scenes	cloud, forest, mountain, plain, sea
large omnivores and herbivores	camel, cattle, chimpanzee, elephant, kangaroo
medium-sized mammals	fox, porcupine, possum, raccoon, skunk
non-insect invertebrates	crab, lobster, snail, spider, worm
people	baby, boy, girl, man, woman
reptiles	crocodile, dinosaur, lizard, snake, turtle
small mammals	hamster, mouse, rabbit, shrew, squirrel
trees	maple, oak, palm, pine, willow
vehicles 1	bicycle, bus, motorcycle, pickup truck, train
vehicles 2	lawn-mower, rocket, streetcar, tank, tractor

임진혁<eeplearning@unist.ac.kr>

Experiment



임진혁<eeplearning@unist.ac.kr>

감사합니다