FL Open-Source Platform Overview

김진수

2022-09-15

Agenda

- Personalization & Benchmark Tool
- (실습) Fedscale
- (실습) FedBench

Federated Learning 문제점

- ‘Advances and Open Problems in Federated Learning’
- 네트워크 토폴로지 및 비동기 통신 문제, 탈중앙 SGD의 로컬 업데이트 문제, 신뢰 문제 등등.....
- 개인화 (Personalization)

→ 연합학습의 자연스럽게 발생하는 특성 ‘이질성(Heterogeneity)’

Method	FedAvg	FedProx	HypCluster	FML	FedMe	LG-FedAvg	FedPer	FedRep	Ditto	pFedMe
Personalization	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓
Clustering	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗
Deep mutual learning	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗
Model splitting	✗	✗	✗	✗	✗	✓	✓	✓	✗	✗
Model update regularization	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓

Personalized FL Benchmark Tool

- 다양한 기술들이 연구를 통해 등장했지만, 명확한 검증들이 진행되지 않음
- 다양한 의문점이 존재
  - Personalization이 과연 현실적인 FL 어플리케이션에서 효과가 있는가?
  - 어떤 Personalization 알고리즘이 최고의 방법인가?
  - Personalization FL과 Standard FL에는 확실한 성능 차이가 있는가?
  - Personalization은 환경 셋팅에 영향을 받지 않는가?
  - …….

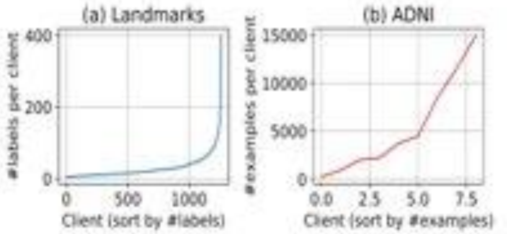
→ 의문점들을 해결할 수 있는 Benchmark Tool 개발(Motley, FedBench)

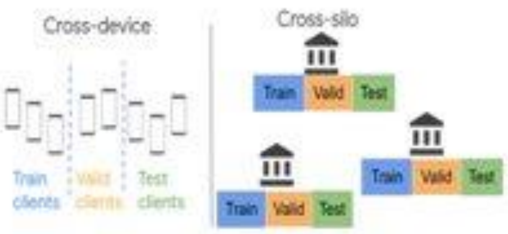
Motley: Benchmarking Heterogeneity and Personalization in Federated Learning

- Cross Device vs Cross Silo 환경의 큰 분류에서 검증 진행
- 사용된 데이터 역시 각 환경에 따라 다른 데이터 사용
- 검증 결과에 따라 Observation 제시

Methods	Dataset Details			
	Dataset	*Task and Model*	*Clients*	*Pts/Client*
Cross-Device FL				
Local training	EMNIST	Image C; CNN	3400	198±89
FedAvg+Fine-tuning	StackOverflow	NWP; LSTM	380k	397±1279
HypCluster/IFCA	Landmarks	Image C; MobileNetV2	1262	130±199
	TedMulti-EnEs	NWP; Transformer	4184	113±56
Cross-Silo FL				
Local training	ADNI	Image R; CNN	9	5405±4822
FedAvg+Fine-tuning	Vehicle	Binary C; SVM	23	1900±349
HypCluster/IFCA	School	R: Linear Regression	139	111±56
Multi-task learning				

Algorithm	Metrics	EMNIST	StackOverflow	Landmarks	TedMulti
Local training	Per-client acc	0.936±.23	0.062±.03	0.215±.17	0.056±.02
FedAvg + Fine-tuning (FT)	Per-client acc before FT	0.852±.08	0.269±.03	0.546±.17	0.160±.04
	Per-client acc after FT	0.989±.02	0.282±.03	0.973±.04	0.162±.04
	% clients "hurt" after FT	0.4%	14%	0.3%	40%
	Best if FT last layer	Y	N	Y	N
	***Practical concerns**: difficult to tune hyperparameters due to local data scarcity; clients may be hurt by FT; sensitive to distribution shift (see Section 4.1)*				
HypCluster / IFCA	Per-client acc	0.897±.07	0.273±.03	0.555±.15	0.163±.04
	No. tuned clusters (k)	2	2	2	2
	% clients largest cluster	53.8%	85.1%	93.1%	54.7%
	Warmstart from FedAvg	Y	Y	Y	Y
	Per-client acc by ensembling k FedAvg models	0.860±.08	0.271±.03	0.564±.16	0.163±.04
	***Practical concerns**: difficult to train due to mode collapse; high communication cost per round; difficult to interpret the learned clusters (see Section 4.2)*				





FedBench: An Empirical study of Personalized Federated Learning

- Performance Comparison
  - Accuracy
  - Convergence speed
  - Training time
  - Communications traffic

- Impact of Experimental Settings on Accuracy
  - Impact of the number of clients
  - Impact of the total number of data samples
  - Impact of the degree of data heterogeneity

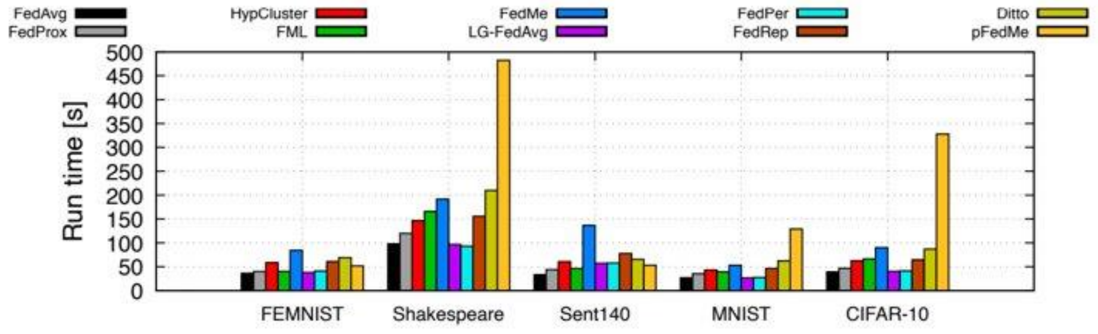


Figure 2: Training time per global communication round.

Table 4: Communication traffic: the number of model parameters communicated between the server and the clients per round.

	FEMNIST		Shakespeare		Sent140		MNIST		CIFAR-10	
FedAvg	2413180	(1×)	1645140	(1×)	161344	(1×)	2399764	(1×)	19870868	(1×)
FedProx	2413180	(1×)	1645140	(1×)	161344	(1×)	2399764	(1×)	19870868	(1×)
HypCluster	3619770	(1.5×)	2467710	(1.5×)	242016	(1.5×)	3599646	(1.5×)	29806302	(1.5×)
FML	2413180	(1×)	1645140	(1×)	161344	(1×)	2399764	(1×)	19870868	(1×)
FedMe	6032950	(2.5×)	4112850	(2.5×)	403360	(2.5×)	5999410	(2.5×)	49677170	(2.5×)
LG-FedAvg	15996	(0.007×)	46260	(0.028×)	25644	(0.159×)	2580	(0.001×)	1060884	(0.053×)
FedPer	2397184	(0.993×)	1598880	(0.972×)	161300	(1×)	2397184	(0.999×)	18809984	(0.947×)
FedRep	2397184	(0.993×)	1598880	(0.972×)	161300	(1×)	2397184	(0.999×)	18809984	(0.947×)
Ditto	2413180	(1×)	1645140	(1×)	161344	(1×)	2399764	(1×)	19870868	(1×)
pFedMe	2413180	(1×)	1645140	(1×)	161344	(1×)	2399764	(1×)	19870868	(1×)

Fedscale

- Fedscale 설치
- Cifar 10 예제 실행
- Fedscale 데이터 다운로드

Fedscale

1. Git clone https://github.com/SymbioticLab/FedScale.git
2. # Please replace ~/.bashrc with ~/.bash_profile for MacOS
3. FEDSCALE_HOME=$(pwd)
4. echo export FEDSCALE_HOME=$(pwd) >> ~/.bashrc
5. echo alias fedscale=\'bash $FEDSCALE_HOME/fedscale.sh\' >> ~/.bashrc
6. conda init bash
7. . ~/.bashrc
8. vi ~/.bash_profile

Fedscale

1. vi environment.yml
2. #   - torch_baidu_ctc==0.3.0

```
#     - torch_baidu_ctc==0.3.0
    - tensorboardX==2.1
    - python-levenshtein==0.12.0
    - pandas==1.1.0
    - PyYAML
    - sox==1.3.7
    - grpcio==1.40.0
    - gym
    - jupyter
```

Fedscale

1. conda env create -f environment.yml
2. conda activate fedscale
3. pip install -e .

Fedscale

1. Fedscale/example/notebook
2. pip install nbconvert
3. jupyter nbconvert *.ipynb --to script

Fedscale

```
import torch
import logging
import math
from torch.autograd import Variable
import numpy as np

import sys, os

from fedscale.core.execution.client import Client
from fedscale.core.execution.executor import Executor
from fedscale.core.logger.execution import args
### On CPU
args.data_dir = './cifar10/'
args.use_cuda = "False"

Demo_Executor = Executor(args)
Demo_Executor.run()
```

Fedscale

1. ./benchmark/dataset/download.sh
2. bash download.sh download feminist(*wget 필요)

FedBench

- FedBench
- 데이터 다운로드
- Femnist 예제 실행

FedBench

1. Git clone https://github.com/OnizukaLab/FedBench.git
2. conda create –n fedbench python==3.7
3. pip install -r requirements.txt

FedBench

1. https://drive.google.com/file/d/1NfmKUFeDogD6DlXkbyhbXI197F3ZfZ02/view
2. 다운로드 받은 데이터를 ./data로 이동

FedBench

1. Leaf 데이터 다운로드
2. Git clone https://github.com/TalwalkarLab/leaf.git
3. cd /data/sent140
4. ./preprocess.sh -s niid --sf 1.0 -k 50 -tf 0.8 -t sample

FedBench

1. cd ./code
2. jupyter notebook
3. 실험 및 테스트 진행
4. 데이터셋을 찾지 못하면 /utils/데이터셋.py/get_data() 경로 조정