Towards Federated Unsupervised Representation Learning

BRAM VAN BERLO, AAQIB SAEED AND TANIR OZCELEBI

System Architecture and Networking Research Group

EdgeSys 2020

Current Situation

Labeled data for model training

- Extensive number centrally aggregated

Ever-increasing number of IoT devices

- Data hard to aggregate
- Active learning is undesirable

Privacy concerns

- GDPR

Supervised architecture	# of labeled training data used
ResNet	1.28M images
Deep Speech 2	English: 12K hours
VGG	1.3M images

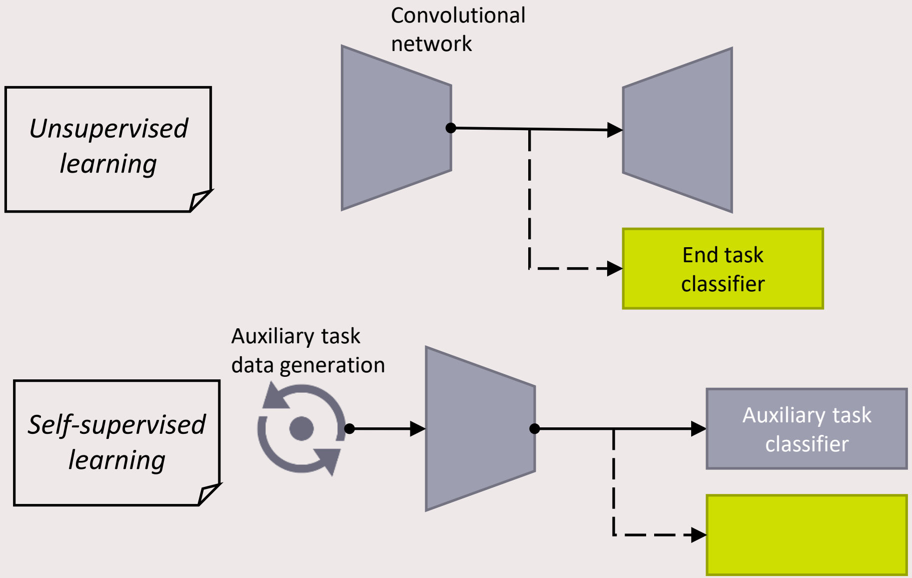
Unsupervised Representation Learning

Methodology

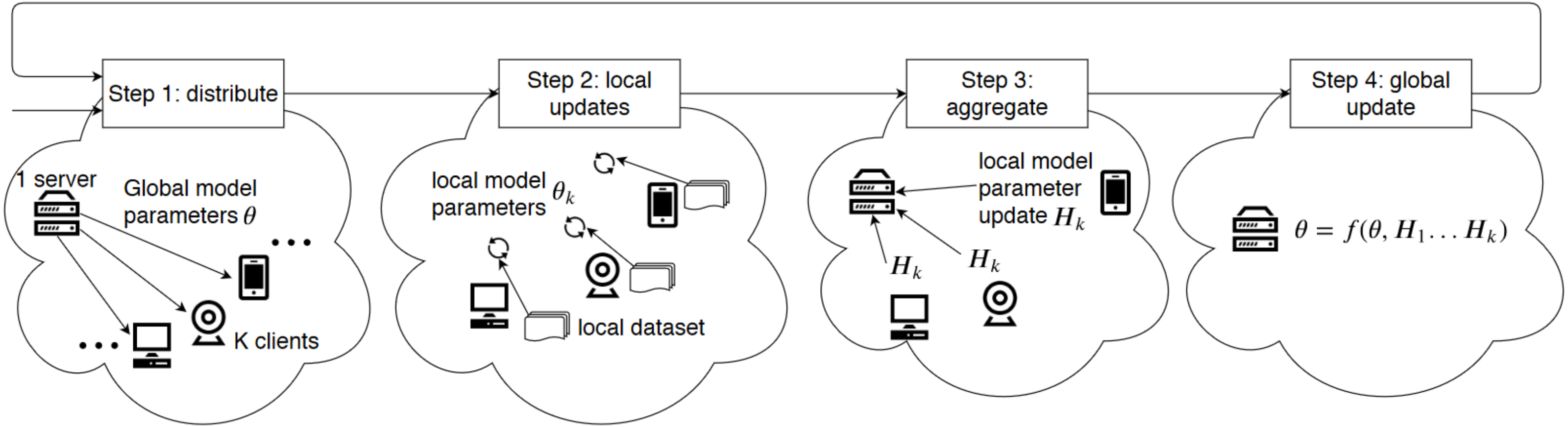
- Pre-train with unlabeled data
- Extract high-level features from labeled data
- Train down-stream task with labeled data

Existing literature

- Limited to centralized context



Federated Machine Learning

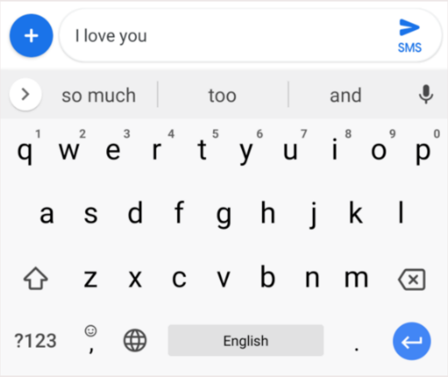


Communication round overview

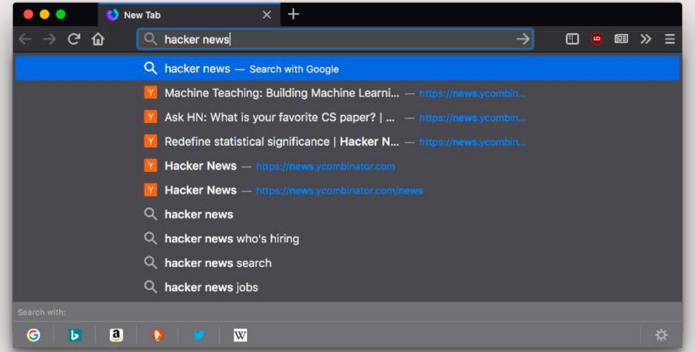
Federated Machine Learning

Existing implementations

- Implicitly labeled data from user interaction



Next word prediction

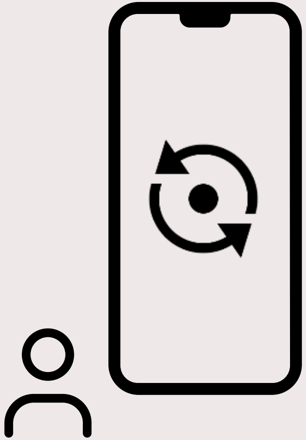


Bookmark suggestion optimization

Federated and Self-supervised Learning

Auxiliary task data generation

- Automated pre-processing step on devices



Compared to active learning

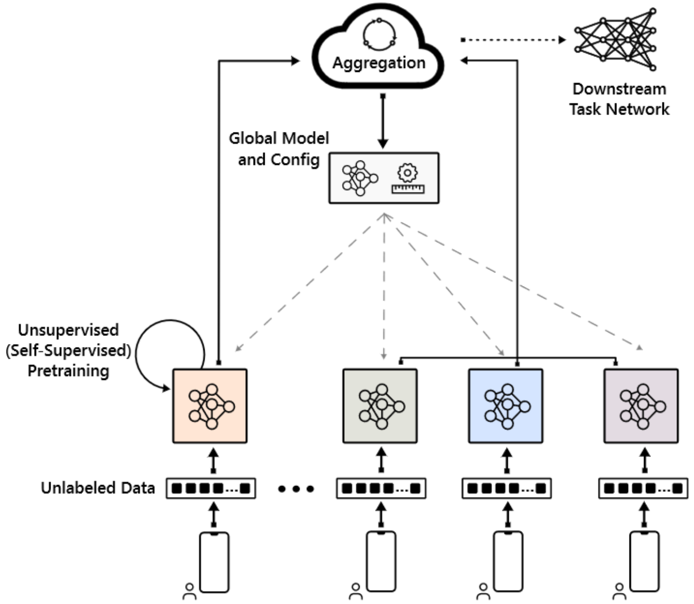
- No human involvement

Contributions

Pre-training models with unlabeled data at IoT devices

Introduction of FURL

Effectiveness demonstration against supervised learning



Effectiveness Demonstration Methodology

Supervised, autoencoder, and self-supervised network

- Encoder-decoder TCN
- Common trunk

Three human activity detection datasets

- HHAR, MobiAct and HAPT

Performance assessment (test data)

- 5-fold user-split cross validation
- Metrics: F1 and Kappa score

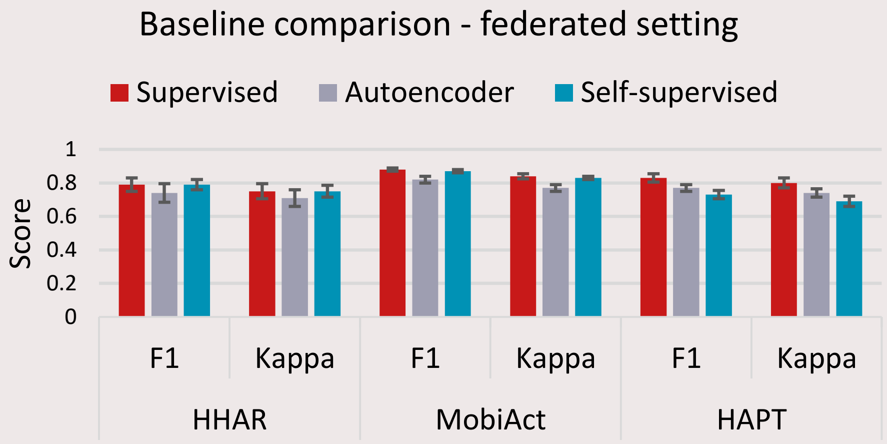
Preliminary findings – baseline comparison

(Un)labeled data share

- Classes
- Generative distribution

Compared to supervised learning: mixed results

- MobiAct, HHAR: on par
- HAPT: worse



Preliminary findings – feature transferability

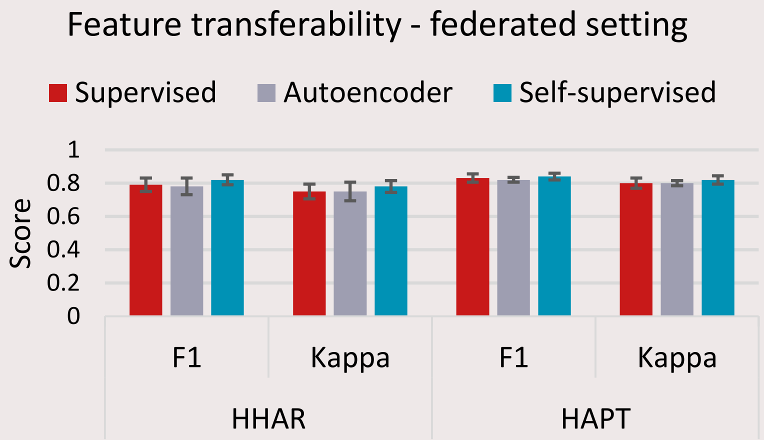
Down-stream end-to-end training on HHAR, HAPT

Pre-training with MobiAct

- More classes
- Bigger generative distribution

Compared to supervised learning

- On average slightly outperforming



Future Work

Experiment extensions

- Non-IID data
- Other application domains

Privacy and security attack handling

Compound model scaling

New modalities and frameworks

Want to collaborate?

Name	Email
Bram van Berlo*	b.r.d.v.berlo@tue.nl
Aaqib Saeed	a.saeed@tue.nl
Tanir Ozcelebi	t.ozcelebi@tue.nl

* Corresponding author