Aalto University
School of Electrical Engineering

ambientintelligence.aalto.fi

Collaborative Machine Learning on Private Data

Le Ngu Nguyen **Stephan Sigg**
Department of Communications and Networking, Aalto University
{le.ngu.nguyen, stephan.sigg}@aalto.fi

Data owners

Data owners are collecting a huge amount of text, images, videos, locations, physiological information, environmental data, et cetera.
They aim to train machine learning models to predict future actions, health status, abnormality, and so on.

Data owners have **privacy** concerns about their sensitive information (e.g. locations and physiological data) and trained models (e.g. model stealing).
In addition, sending data is expensive.

System

Data owners send **all of their raw data** and the system trains a global model using a **centralized** method: data and model are kept by the system.

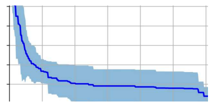
We introduce a model training algorithm in which:
+ Data owners keep model weights w_i and data x_i
+ Data owners only send combination of w_i and x_i
+ Data is vertically-partitioned, i.e. each owner has different attributes for the same set of training samples.
+ No party can access the complete model.

Our Collaborative Model Training Method

Collect data x_i and calculate $w_i x_i$ — *Send $w_i x_i$* → Calculate $\sum w_i x_i$ from all users
Evaluate the model quality
Feedback: positive (model improved) or negative (model not improved)

Update weight $w_i \rightarrow w'_i$ and collect data x'_i ← *Send binary feedback*
+ Positive feedback: continue the same update direction
- Negative feedback: restore the previous value & alternate the update direction

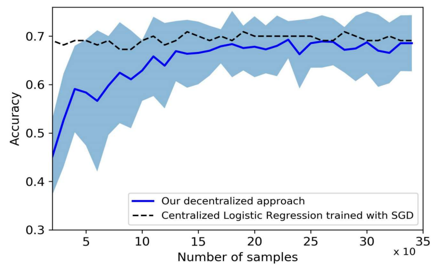
Send $w'_i x'_i$ → Repeat until convergence



Experiments and Results

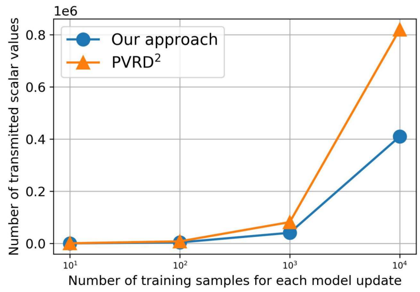
Classification performance of our privacy-aware training procedure

Our algorithm has been experimented with: occupancy monitoring, intrusion detection, phishing detection, vehicle classification, and diabetes diagnosis.
The number of samples is from thousands to hundreds of thousands.
The number of attributes (features) is up to 100.
We achieve competitive accuracy on the test set comparing to the centralized approach.



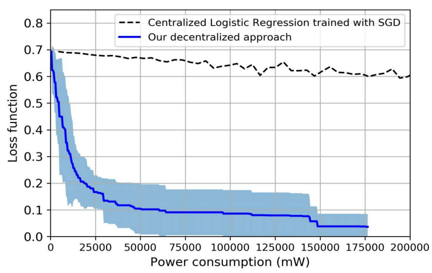
Amount of shared data

We transmit significantly less shared data, comparing to a consensus approach (users exchange data with their neighbours to update the model until convergence)
PVRD²: Pipelined variance-reduced dynamic diffusion (Ying et al., Supervised Learning Under Distributed Features, IEEE Transactions on Signal Processing, 2019.)



Energy efficiency

Our algorithm requires less power to converge.
Thus, it is suitable for resource-constrained settings.



Future work

We will integrate encryption protocols to strengthen our method.
We will experiment our method on resource-constrained devices.

N. Nguyen and S. Sigg, Learning a Classification Model over Vertically-Partitioned Healthcare Data, *IEEE Multimedia Communications – Frontiers*, 2019.