Distributed Private Machine Learning

Abhradeep Guha Thakurta

University of California Santa Cruz
Distributed learning from private data
Distributed machine learning - Setup

Traditional machine learning

\[ \theta : \text{Model / statistic} \]

Data

Distributed machine learning

\[ \theta \]
Learning from private data
Learn new (and frequent) words typed

Phablet
Derp
Photobomb
Woot

Phablet
OMG
Woot
Troll

Prepone
Phablet
awwww
dp
Learning from private data

Predict if a person has Parkinson’s disease

Get measurements from gyroscope, display screen etc.

Model / classifier

YES / NO

Image courtesy: Research kit (Apple)
Assumption: Hidden matrix has some structure (e.g., low-rank)
Local differential privacy \cite{Warner65,EGS03,DMNS06}

Data sample: \( d \)

\[ \mathcal{A}(d) \]

Data sample: \( d' \)

\[ \mathcal{A}(d') \]

Requirement: \( \mathcal{A}(d) \) and \( \mathcal{A}(d') \) should be close in distribution
Local differential privacy [Warner65,EGS03,DMNS06]

$\epsilon$: Privacy parameter (smaller value implies stronger privacy)

Resilient against arbitrary side information

Provably protects against membership attacks
Challenge: Balancing trade-offs

Conflicting goals

Balancing the tradeoff is hard:

• AOL fiasco: *CNBC 101 dumbest moments in business*
• Netflix attack [NS08], Facebook attack [Korolova11], ...
This talk

Conflicting goals

Utility

Privacy

Interaction
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning
Private distributed model selection
Learning from private data

Predict if a person has Parkinson’s disease

θ: Model / classifier

YES / NO
Towards engineering distributed learning systems

Ideal scenario: Complete parallelism

Each device interacts with server independently and only once

$\theta$
Towards engineering distributed learning systems

State of the art [DJW’13]: Completely adaptive interaction

Server must:
• Talk to devices in sequence
• Receive message from each device in order to compute message to next device
This talk [Smith T. Upadhyay’ 17]

Distributed private learning with local differential privacy

New algorithms that use little or no adaptive interaction

Lower bound: Cannot get accurate, general algorithms that use no adaptive interaction
Previous work

Distributed private learning with local differential privacy

Kasivishwanathan \textit{et al.} 2008

Introduced the problem of local private learning

Duchi \textit{et al.} 2013

Tight upper and lower bounds on accuracy

\# of adaptive interaction = \# of devices

Talks to each device only once
Key New Results

Single parameter learning: Minimal error with full parallelism

Diff. private model
\( \theta_{priv} \in \mathbb{R} \)

Translate to...

Diff. private estimate of Median \( \{d_1, d_2, \ldots, d_n\} \)
Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity
Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity
Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity

Diff. private model $\theta_{priv} \in \mathbb{R}^p$

Exponential improvement in the rounds of adaptivity

Still interact with one device only once
Key New Results

Lower bound: Minimal error needs few rounds of adaptivity

$\theta^*$: Best model
Next Steps

Implement the algorithms and evaluate empirically

Deploy the project in practice

Current lower bounds are only for gradient based methods

• Obtaining non-adaptive algorithms will analyzing non-gradient based methods
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning
Private on-device learning
On-device learning with sensitive data

Privacy preserving personalization

- Health analytics
- Language models
- Collaborative filtering

Locally learned models

Model estimates via differentially private channel

Receive model updates

Local learning
On-device learning with sensitive data

Privacy preserving personalization

Health analytics
Language models
Collaborative filtering

Local learning

Model estimates via differentially private channel

Receive model updates

Locally learned models
On-device learning with sensitive data

Local computation
- Diff. private in regards to everyone else’s data

Model estimates via differentially private channel

Receive model updates

Locally learned models
- Diff. private in regards to everyone’s data

Global computation
- Diff. private in regards to everyone’s data

Privacy preserving personalization
Differentially private on-device learning

New results and future direction

Collaborative filtering [Jain T. Thakkar]

First algorithm with formal error guarantee

- Global component: Error covariance

Add noise and send

Average error covariance across all devices
Differentially private on-device learning

New results and future direction

Collaborative filtering [Jain T. Thakkar]

First algorithm with formal error guarantee

- Global component: Error covariance
- Local component: Compute the prediction

[HR12] Hints at trivial error if predictions are public

Next step

Improve on-device machine learning by harnessing global computation
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning
Conflicting goals

- Utility
- Privacy

Interaction