A Differentially Private Stochastic Gradient Descent Algorithm for Multiparty Classification

Arun Rajkumar - Shivani Agarwal

What is the problem

- developing privacy-preserving machine learning algorithm in a distributed multiparty setting
- different parties own different parts of data set
- GOAL: learn a classifier from the entire data set, without revealing any informations about the individual data points it owns
- new differentially private algorithm for the multiparty setting that uses stochastic gradient descent based procedure

How does it address the problem / preliminaries

- Differential Privacy
  - A - randomized algorithm
  - D_1, D_2 datasets that differ on a single element
    - A is ϵ-differentially private for D_1 and D_2,
      if $P[A(D_1)] \leq e^{\epsilon} \times P[A(D_2)]$
- Differential Privacy in Machine Learning: output and objective perturbation
  - Output P. : adds noise to the output of the algorithm, perturbs the output classifier
  - Objective P. : perturbs the objective minimized by the algorithm
- Multiparty Classification
  - Learning a classifier that preserves the privacy of each party

What results were presented

- Basic Idea
  - minimize the overall multiparty objective by running a gradient descent algorithm
  - there is not taken any privacy consideration
- Privacy consideration
  - Attempt 1: adding $\boldsymbol{\rho}_t^k$ noise to the gradient (t - time, k - k^{th} party) - not enough, the third party can converge to the true minimizer w*
  - Attempt 2: adding another noise vector η^k - sampled only once by each party P_k

What do I think was good and bad about this paper

- Good
  - the intention / the will
  - attempts / multi-level perturbing
  - simple and easy, but effective algorithms
- Bad
  - short explanation
  - too low proving
  - befogging