Assisted Learning: Cooperative AI with Autonomy

Jiaying Zhou
Department of Statistics
University of Minnesota, Twin Cities
zhou1054@umn.edu

Author: Jiaying Zhou, Xun Xian, Na Li, Jie Ding

Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Assistance between Entities

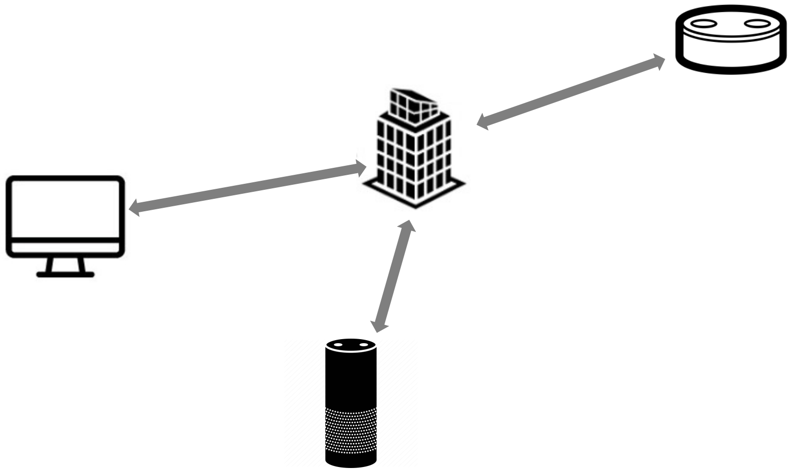
- Many entities, each possesses private data, model
- Assistance is possible via some coordinates.
- Examples
  - Two autonomous car data in different locations
  - The same group of mobile users beheld by different entities

Assistance between Entities

- The challenge of centralizing data?
  - Privacy issue: data/ model are sensitive, cannot be shared
  - Transmission cost: large features, e.g., video dataset
- How can each entity leverage the data and computation resources without leaking exclusive data/ model information?
- Related work: Federated Learning

Federated Learning

- Related work: Federated Learning [1-3]
  - Goal: leverage resources of edge devices to achieve a **global objective**
    - Example: many mobile users -> user interest
  - Method: learn a global model using the **averaging** of locally learned model parameters
  - Characteristics: no data sharing, **a central server, a public model and objective**



Recent trend:
entities have heterogeneous model[4]; autonomy

[1] R. Shokri and V. Shmatikov, “Privacy-preserving deep learning,” in Proc. CCS. ACM, 2015.
[2] J. Konecny, H. B. McMahan, F. X. Yu, P. Richtarik, A. T. Suresh, and D. Bacon, “Federated learning: strategies for improving communication efficiency,” arXiv preprint arXiv:1610.05492, 2016.
[3] H. B. McMahan, E. Moore, D. Ramage, S. Hampson et al., “Communication-efficient learning of deep networks from decentralized data,” arXiv preprint arXiv:1602.05629, 2016.
[4] E. Diao, J. Ding, V. Tarokh, “HeteroFL: Computation and Communication Efficient Federated Learning for Heterogeneous Clients,” arXiv preprint arXiv:2010.01264l.

Assisted Learning

- Goal:
  - To allow entities to improve each other's learning capability with high autonomy, entities have heterogeneous model/data without transmitting private information.

- Characteristics:
  - Privacy: No one will share private data/model
  - Communication: as few as possible (often within 20) particular for large organization
  - Autonomy: There need not be a central controller to organize the learning process

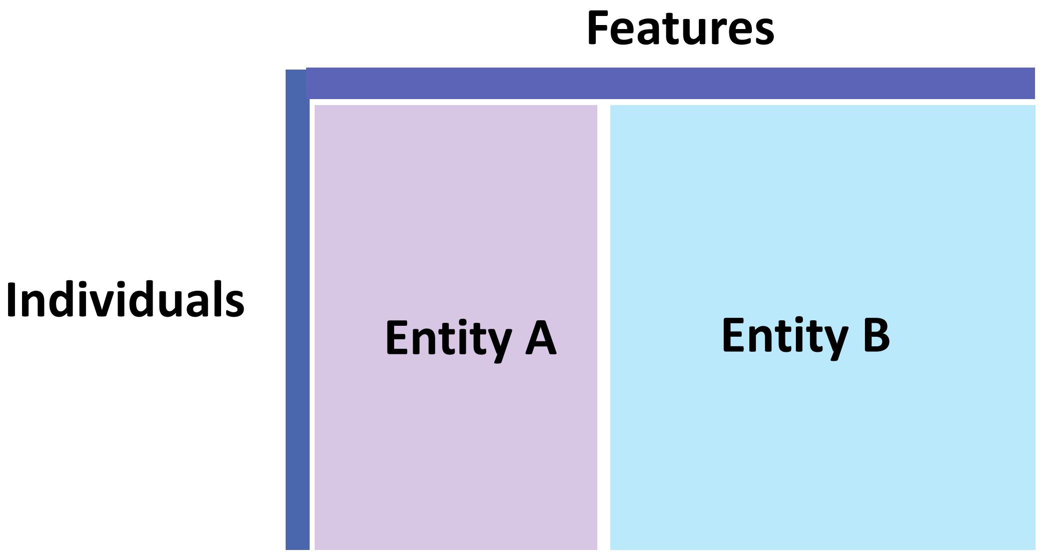
Assisted Learning Methods

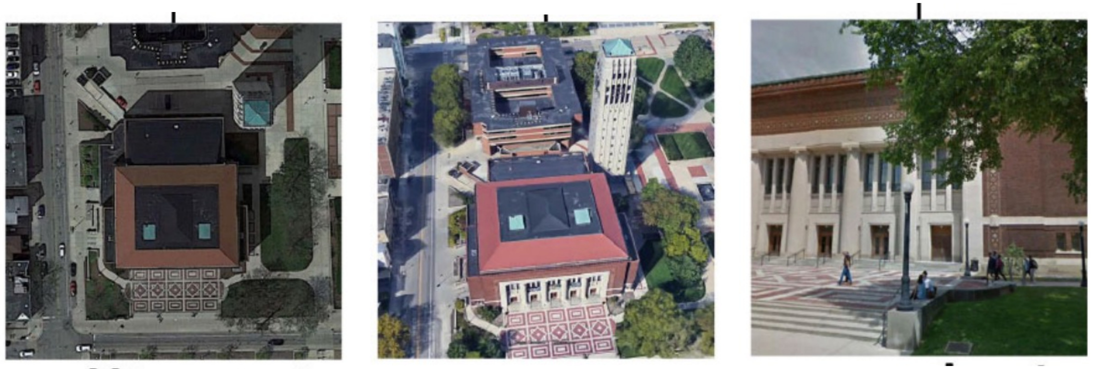
- What information to exchange in Assisted Learning?
  - There is no global model! - No unified method
  - Existing
    - **Fitted residuals [1]**
    - **Scope: regression**
  - (Proposed) **ASCII**: broader scenario (any supervised setting/model)
    - **Ignorance scores**
    - **Scope: classification (main)**/regression

[1] Xian, Xun, et al. "Assisted Learning: A Framework for Multiple Organizations' Learning." *arXiv e-prints* (2020): arXiv-2004.

Setting: partial features, shared individuals

- Vertically splitting setting: partial features, shared individuals (IDs)
  - the same cohort of mobile users observed by different entities
  - the residents at Minnesota and their financial/healthcare providers
  - photos taken from different angles.





Satellite view **Drone** view **Ground** view

Example: **Object detection coordinated by 3 entities**

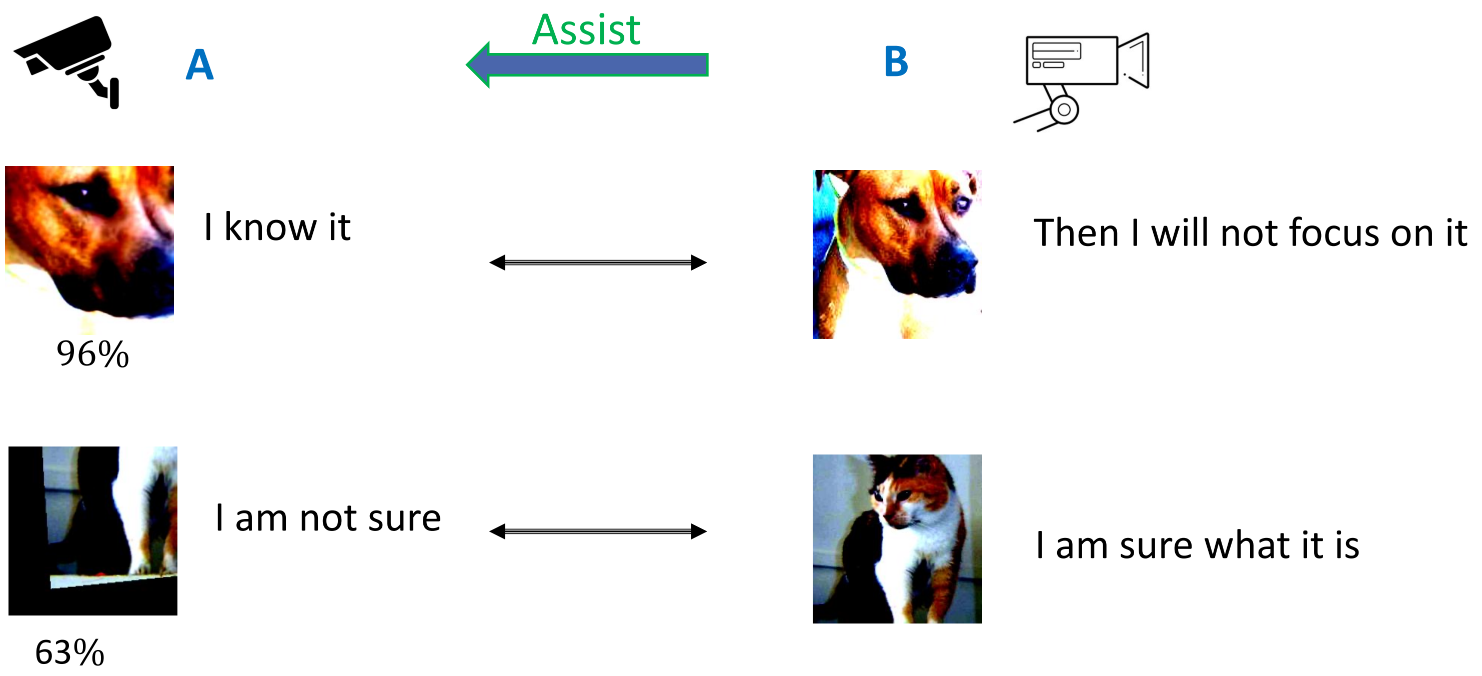
Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Key Idea: ASCII at training stage

- Modeling and decision making without sharing proprietary info?

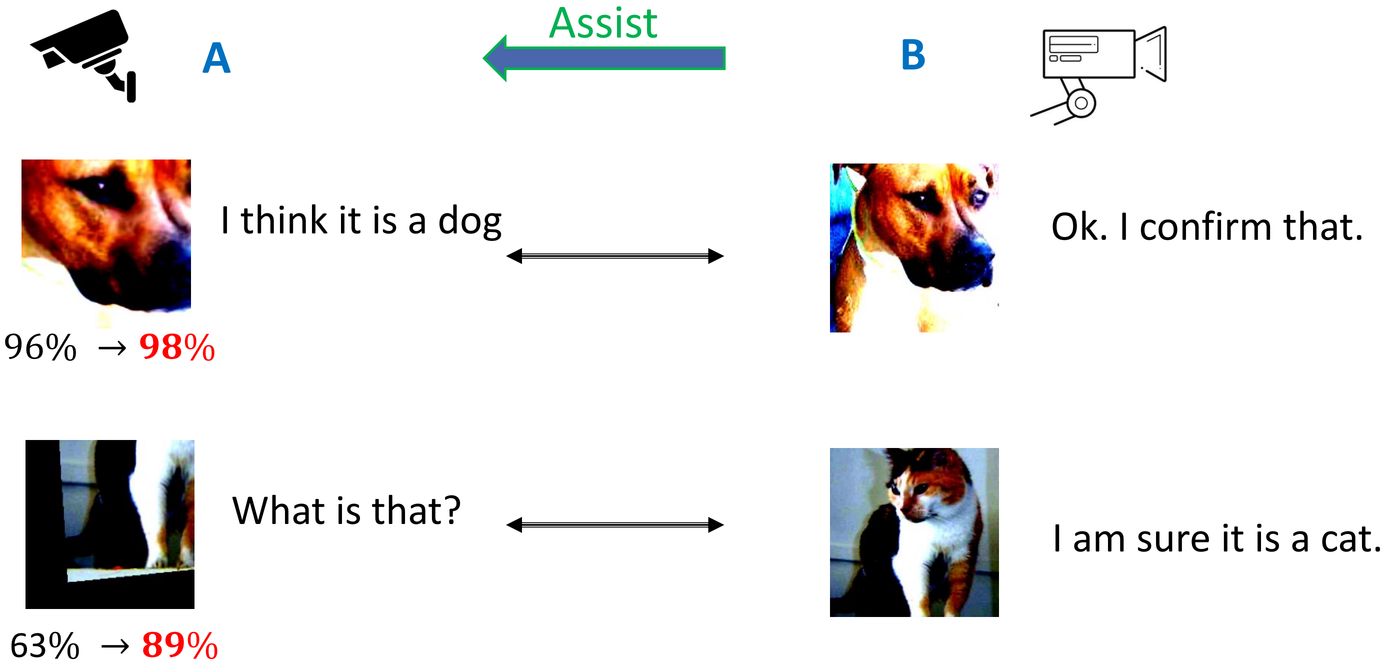
Assisted Training Stage



Key Idea: ASCII at prediction stage

- Modeling and decision making without sharing proprietary info?

Assisted Prediction Stage



Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Problem Formulation

- Technical formulation
  - **Key idea: turn an unrealistic optimization problem that depends on centralized data into one that can be operated in a decentralized fashion**

- Additive nonparametric model in hindsight

Number of assistance rounds

$$\mathcal{F}_T = \left\{ \mathbf{f}_T \left(\mathbf{x}_i^{(1)}, \ldots, \mathbf{x}_i^{(M)} \right) = \sum_{t=1}^{T} \sum_{m=1}^{M} \alpha_t^{(m)} \mathbf{g}_t^{(m)} (\mathbf{x}_i^{(m)}) \right\}$$

Supervised function class

Private data of M agents

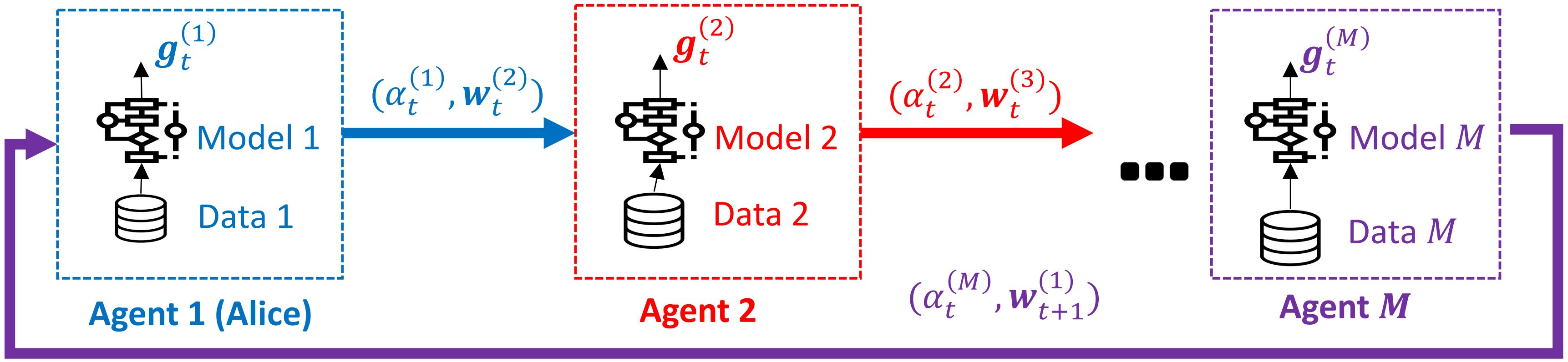
Private model of each agent

- Exponential loss $\mathfrak{L} : (\mathbf{y}, \mathbf{f}_T) \mapsto e^{-\frac{1}{K}\mathbf{y}^{\mathrm{T}}\mathbf{f}_T}$

Label (encoded)

Predicted label (pre-softmax)

ASCII: Pipeline



ASCII: Demonstration of the algorithmic update in the presence of M agents.

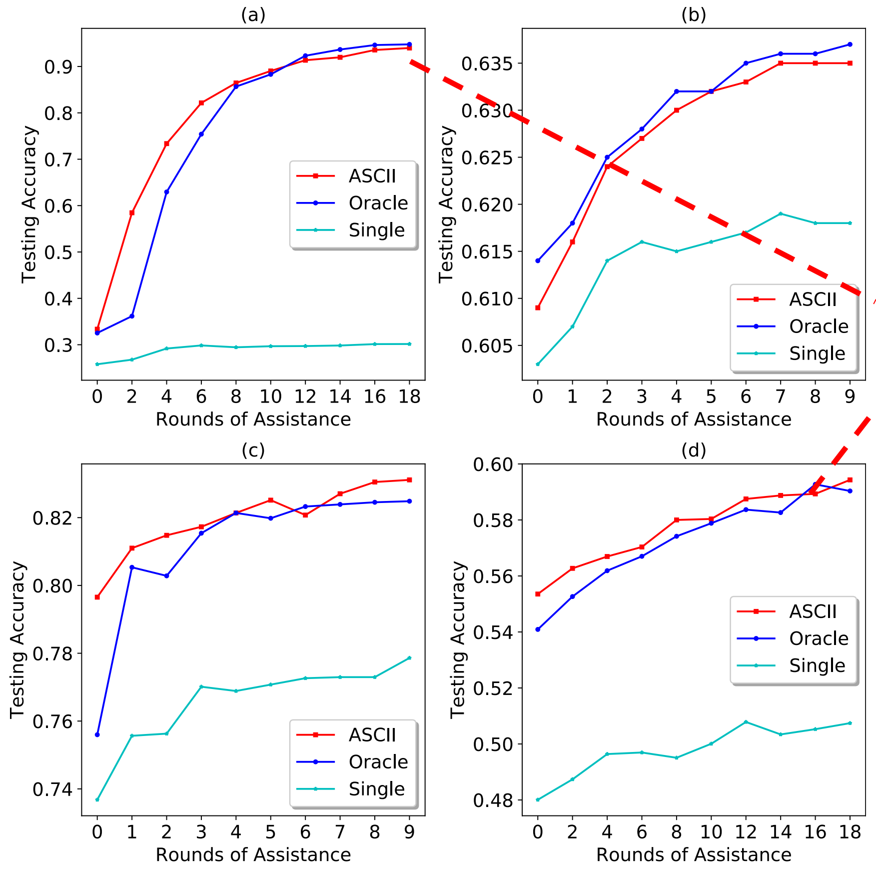
- When B assists A, transfer: **sample ignorance scores**
- Ignorance score: value between 0 and 1; smaller value -> confident
- Predictive additive model: $\mathcal{F}_T = \left\{ \mathbf{f}_T \left(\mathbf{x}_i^{(1)}, \ldots, \mathbf{x}_i^{(M)} \right) = \sum_{t=1}^{T} \sum_{m=1}^{M} \alpha_t^{(m)} \mathbf{g}_t^{(m)} (\mathbf{x}_i^{(m)}) \right\}$

Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Experiment: Improve Accuracy

- Data: Synthetic Gaussian Blobs, MIMIC 3, QSAR, Wine
- Oracle: Unrealistic pulled data
- Single: Agent A
- ASCII is always better than non-assistance!



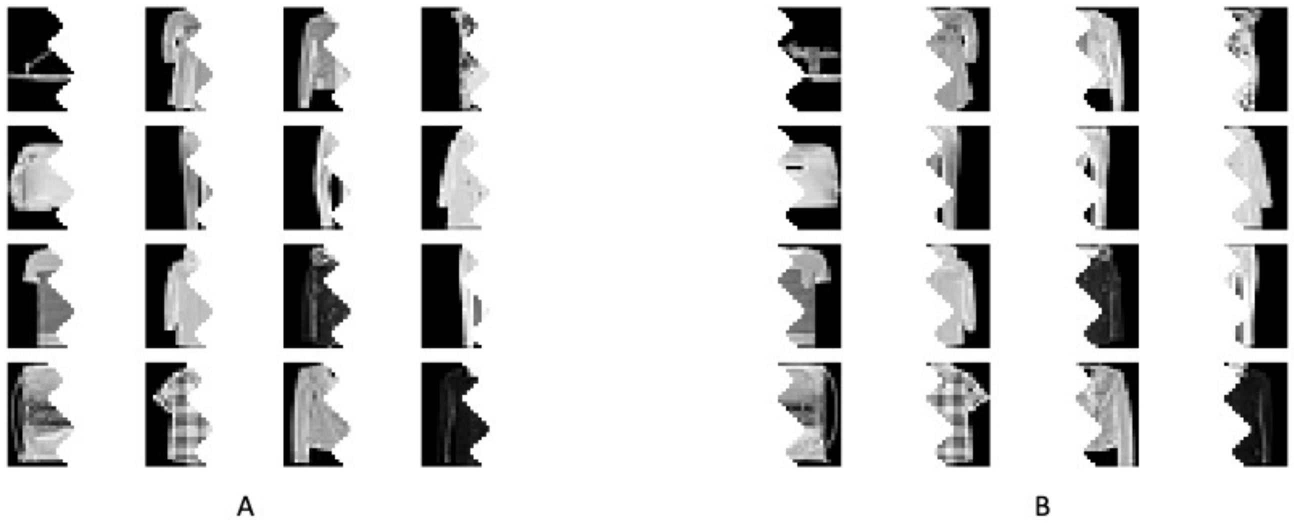
Performs like they pull together!

Figure: ASCII improves prediction accuracy

Experiment: Reduce Transmission Cost

Noisy Gaussian Blobs Data: data is generated from five features, 195 noisy features, two agents.

Fashion-MNIST

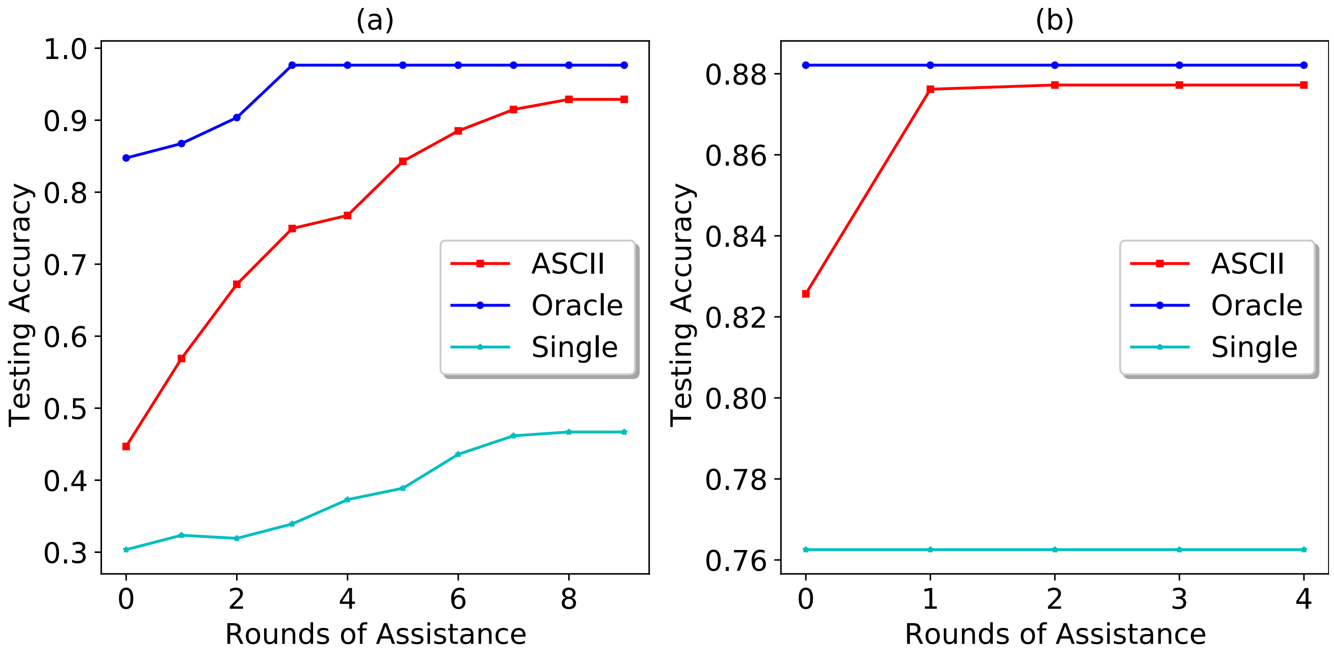


A holds 1/2 of the picture

B holds another 1/2

B assists A

Experiment: Reduce Transmission Cost



Compared with Oracle: Save transmission cost, protect privacy

Compared with non-assistance: improve performance

Figure: Out-sample predictive accuracy for the datasets of (a) Noisy Gaussian Blob data, (b) Fashion-MNIST. The transmission costs of (a) and (b) are improved by around 10 and 195 times compared to the oracle approach.

Outline

- Assisted Learning
- ASCII
  - Key Idea
  - Algorithm
  - Experiments
- Summary

Summary

- Assisted learning
  - Assistance between entities is important
  - Assistance without sharing private data/ model …
- ASCII algorithm
  - Efficient, usually converges in less than 20 rounds
  - Private, entities keep local model/ data private
  - Save transmission cost, especially in large dataset, e.g., video dataset
  - Autonomous, there is no central controller

Thank You!