

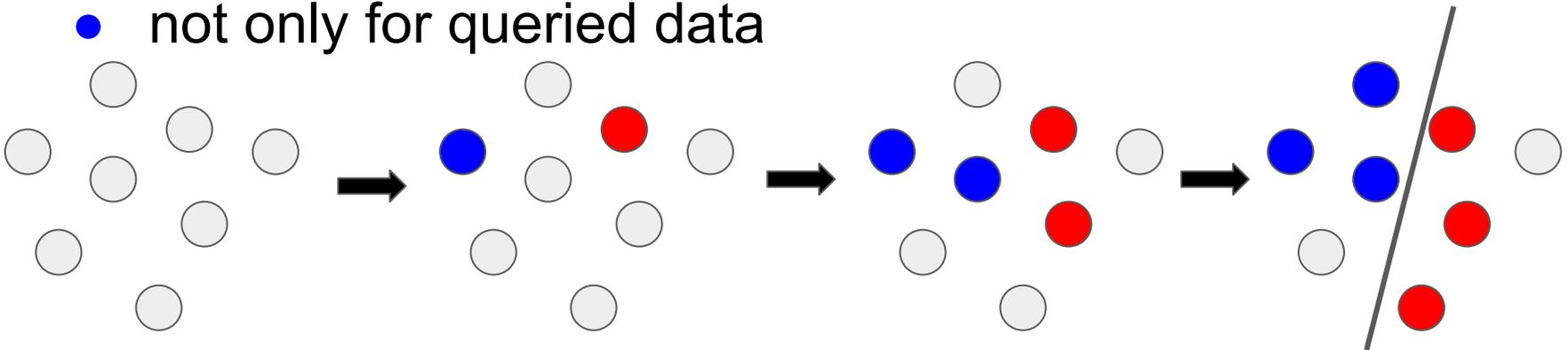
Active Learning Using Discrepancy

Zhenghang Cui, Issei Sato
The University of Tokyo / RIKEN AIP Center

RealML @ ICML2020
July 18, 2020

(Pool-based) Active Learning

- Given:
 - a pool of unlabeled data
 - a labeling oracle
- Decide for each step:
 - query batch of unlabeled data
- Goal: a classifier for all data
 - not only for queried data

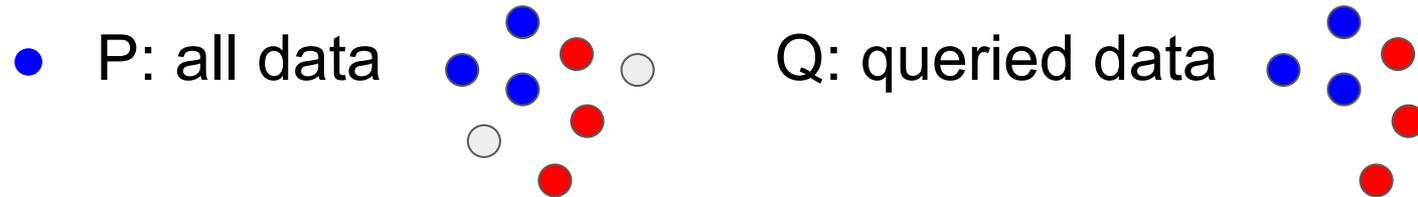


Deep Batch Active Learning

- New requirements:
 - Using deep neural network classifiers.
 - Large query batch for each step.
- Traditional active learning theory can hardly help.
- Selecting batch with **diversity** and **representativity**:
 - [Sener and Savarese, ICLR, 2018] on corsets
 - [Ash et al., ICLR, 2020] on loss gradients
 -
- **Unified** method
 - [Shui et al., AISTATS, 2020] on the Wasserstein Distance

Proposed Method using Discrepancy

- Consider two distributions:



- With two functions f and f' , we minimize the error bound on \mathbb{P} established using

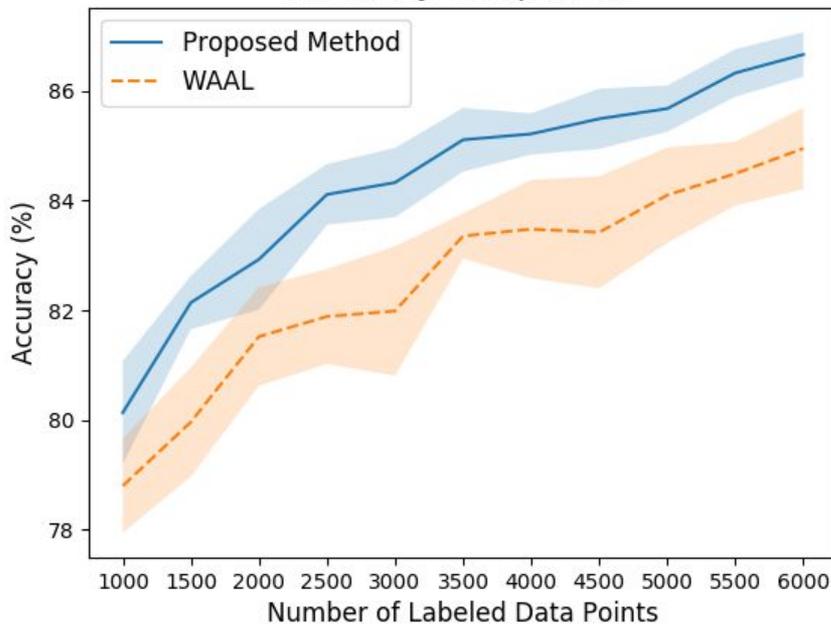
$$\text{disc}_{f, f'}(\mathbb{P}, \mathbb{Q}) = \left| R_{\mathbb{P}}^{\ell}(f, f') - R_{\mathbb{Q}}^{\ell}(f, f') \right|$$

- Lemma 3 (informal):
 - Above discrepancy is upper bounded by the Wasserstein Distance.

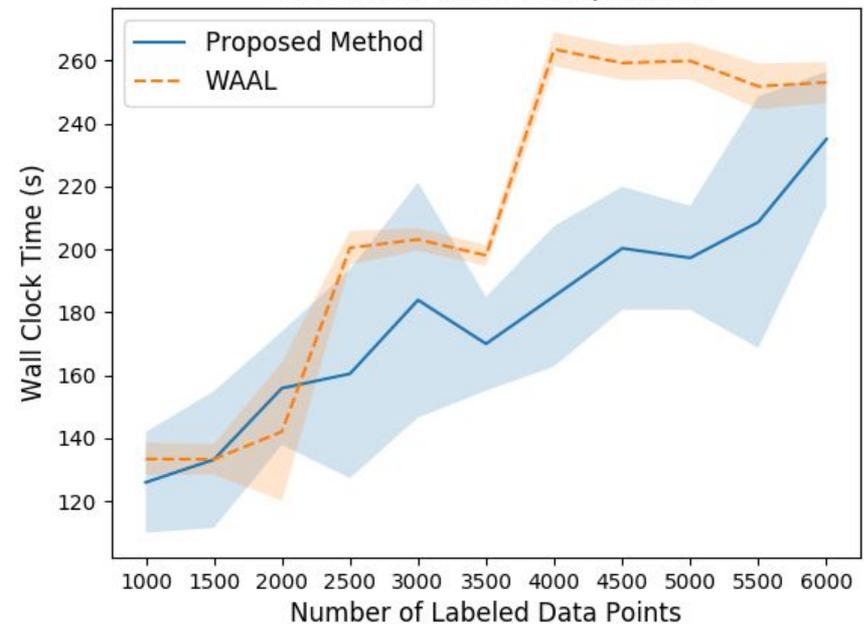
Experimental Results

- Fashion-MNIST Dataset [Xiao et al. 2017]
- 10 repetition under same setting as SOTA 'WAAL' [Shui et al., AISTATS, 2020].

Accuracy Comparison



Execution Time Comparison



Thank you!