Problem setting
- Problem: Multiparty multiclass classification
- Example: Flu detection

But, the types of flu diverse geographically, the distribution of patients records collected by a hospital in California is different from Florida. Good local models are built:

- Local model in California detects [1,2,3] types
- Local model in Florida detects [3,4] types

The patients’ records are confidential. Can we smartly reuse the local models to learn the global problem, instead of building a model on merged local datasets?

Contribution
Q: How to measure the global behavior of multiple models?  
A: Multiparty multiclass margin (MPMC-margin)

Q: How to optimize the global behavior?  
A: The HMR method, which maximizes MPMC-margin, by modifying local models, without merging local datasets.

MPMC-margin
- The multiparty multiclass margin (MPMC-margin) on the local predictors set $H=\{h_1, \ldots, h_k\}$ at a labeled example $(x,y)$ is defined as:
  $$
  \rho_H(x,y) = \max_{i \in [k]} h_i(x,y) - \max_{j \in [k]} h_j(x,y'),
  $$
  
  where $y \in \mathcal{Y}_i, y' \in \mathcal{Y}_j \setminus \{y\}$.

  
  Non-positive MPMC-margin causes wrong prediction, so we want to maximize it.

Heterogeneous model reuse method
An iterative method exchanges $T$ examples and maximizes MPMC-margin on “unobserved” merged global dataset.

Experiments
- Toy example on LR/SVM/GBDT
- Heterogeneous learning models
- LR: green, yellow
- SVM: green, magenta
- GBDT: magenta, orange
- Exchanged 20 examples
- Nearly perfect performance

- Benchmarking on fashion-MNIST on various data partitions

- Multi-lingual handwriting recognition
  - 6 different structured CNNs trained locally on hiragana, katakana, kanji, devanagari, hangul and English letters
  - 1600+ classes, 94.32% global accuracy
  - Only exchanged 300 out of 420k examples

Contact: wuzx@lamda.nju.edu.cn, wuzx.gm@gmail.com
Code: https://github.com/YuriWu/HMR