Binary Classification from Positive-Confidence Data

Takashi Ishida[1,2] **Gang Niu**[2] **Masashi Sugiyama**[2,1]

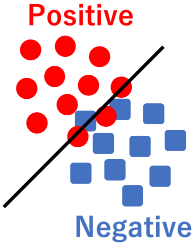
[1] The University of Tokyo
[2] RIKEN

NeurIPS 2018, Canada, December 6th, 2018

Introduction

- Ordinary classification:

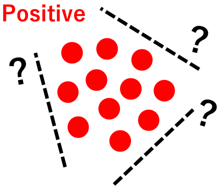
  Learn a binary classifier with *both* **positive** and **negative** training data.



- Research question:

  Can we learn a binary classifier from *only* **positive** data?

  Without any **negative** data, or even **unlabeled** data?

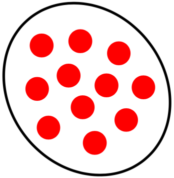


Related Works

<u>One-class classification</u>

- **Describe** the positive class by clustering-related methods
- Cannot tune hyper-parameters for maximizing the generalization ability

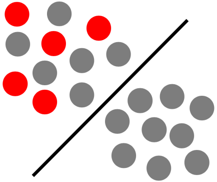
Positive



<u>Positive-unlabeled classification</u>

- *Additional* **unlabeled** data from marginal distribution
- Assumes that the class prior $p(y)$ is known

Positive



Unlabeled

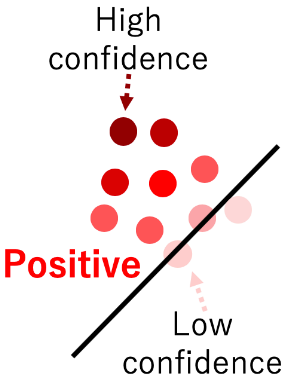
Main Idea

Equip positive data with confidence:

- Example: 95% DOG (5% WOLF)

Main message of the paper:

- If you can equip positive data with confidence (positive-confidence), you can learn a binary classifier with **optimal convergence rate**!
- Positive-confidence includes the information of the negative distribution $\rightarrow$ allows us to discriminate between positive/negative classes.
- *Positive-confidence (Pconf) classification*.



Empirical Risk Minimization (ERM) in Binary Classification

- Train a binary classifier $g(\boldsymbol{x})$ so that the classification risk $R(g)$ is minimized:

$$R(g) = \mathbb{E}_{p(\boldsymbol{x},y)}[\ell(yg(\boldsymbol{x}))]$$

  - Input $\boldsymbol{x} \in \mathbb{R}^d$ and its class label $y \in \{\pm 1\}$ follows $p(\boldsymbol{x}, y)$.
  - $\mathbb{E}_{p(\boldsymbol{x},y)}$ denotes expectation over $p(\boldsymbol{x}, y)$, $\ell(z)$ is loss function

Setting of Pconf Classification

- **Issue**: We do not have data from $p(\boldsymbol{x}, y)$!
- Only have positives equipped with *confidence*:$\mathcal{X} := \{\boldsymbol{x}_i, r_i\}_{i=1}^n$
  - $\boldsymbol{x}_i$ is positive data drawn independently from $p(\boldsymbol{x}|y = +1)$.
  - r_i is the positive-confidence given by $r_i = p(y = +1|\boldsymbol{x}_i)$.

Main Contribution

Theorem

Classification risk can be expressed as

$$R(g) = p(y=+1) \cdot \mathbb{E}_{p(\boldsymbol{x}|y=+1)}\left[\ell(g(\boldsymbol{x})) + \frac{1-r(\boldsymbol{x})}{r(\boldsymbol{x})}\ell(-g(\boldsymbol{x}))\right],$$

if we have $r(\boldsymbol{x}) \neq 0$ *for all* $\boldsymbol{x}$ *sampled from* $p(\boldsymbol{x})$*, where* $r(\boldsymbol{x}) = p(y=+1|\boldsymbol{x})$.

- Leads to a theoretically-grounded way of objective function design for Pconf classification!

Summary of the Paper

- Propose a novel setting/algorithm for classification from only positive data equipped with confidence
- Establish an ERM counterpart to binary classification for this setting
- Confirm the potential problem of a naive objective and prove the optimal convergence rate of the proposed ERM objective
- Validate the proposed ERM objective with broad experiments

Our poster @ Room 210 & 230 AB #97