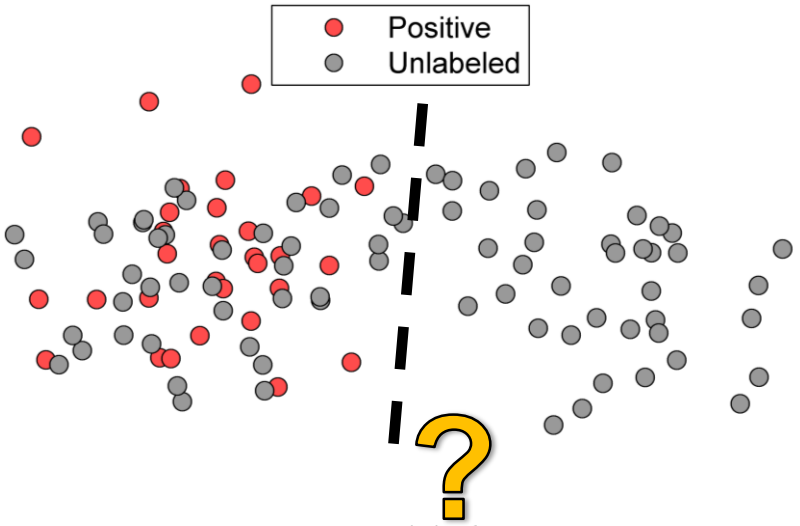
T-31 Learning from Positive and Unlabeled Data 1: Classifier Training and Theoretical Analysis

Marthinus C. du Plessis (UTokyo), Gang Niu (Baidu Inc.) and Masashi Sugiyama (UTokyo)

Task: Train a classifier using only *positive and unlabeled* data

Given *an estimate of* π, we:

1. Show how to train a classifier
2. Show the effect of estimation error of π
3. Give generalization bounds



$$\mathcal{X} := \{\boldsymbol{x}_i\}_{i=1}^{n} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x}|y=1)$$

$$\mathcal{X}' := \{\boldsymbol{x}'_i\}_{i=1}^{n'} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x})$$

$$p(\boldsymbol{x}) = \pi p(\boldsymbol{x}|y=1) + (1-\pi)p(\boldsymbol{x}|y=-1)$$

$$\pi = p(y=1) \quad \textit{Class prior}$$