PU Classification

$$
\begin{aligned}
R(g) &= \mathbb{E}_{(X,Y)\sim p(x,y)}\left[\ell\big(Yg(X)\big)\right] \\
&= \pi_{\mathrm{p}}\mathbb{E}_{\mathrm{p}}\left[\ell\big(g(X)\big)\right] + \pi_{\mathrm{n}}\mathbb{E}_{\mathrm{n}}\left[\ell\big(-g(X)\big)\right] \\
&= \pi_{\mathrm{p}}\mathbb{E}_{\mathrm{p}}\left[\ell\big(g(X)\big)\right] + \mathbb{E}_{X}\left[\ell\big(-g(X)\big)\right] - \pi_{\mathrm{p}}\mathbb{E}_{\mathrm{p}}\left[\ell\big(-g(X)\big)\right]
\end{aligned}
$$

Unbiased PN estimation

PU SSL

$$R^{\gamma}_{\mathrm{PU+PN}}(f) = \gamma R_{\mathrm{PU}}(f) + (1-\gamma) R_{\mathrm{PN}}(f) \quad 0 \le \gamma \le 1$$

PU-based SSL 中存在的一些问题:
- 半监督学习中，更常见的是one sample问题。
- 在 PU-based SSL 中需要明确知道正负类的先验
- 参数 γ 挑选的问题。

$$\begin{aligned}
R_{PU}(g) &= \pi_P R_P^+(g) + \left\{ R_U^-(g) - \pi_P R_P^-(g) \right\} \\
&= \left\{ \pi_P R_P^+(g) + \pi_N R_N^-(g) \right\} + R_U^-(g) - \left\{ \pi_P R_P^-(g) + \pi_N R_N^-(g) \right\} \\
&= R_{PN}(g) + R_U^-(g) - R_{PN}^-(g) \\
&\approx R_L(g) + \left\{ R_U^-(g) - R_L^-(g) \right\}
\end{aligned}$$

需满足独立同分布假设

Supervised　Semi-Supervised

解释:
- 第一项是监督学习的风险
- 第二项是在未标记数据和标记数据上风险的差。如果已标记数据数据满足独立同分布假设，这一项的值应该是小的。

PUAL

g: 当前分类器
$g^{+(x_k,y_k)}$: 加入样本(x_k,y_k)得到的分类器

$$x^* = \arg\min_{x_k \in X_U} R(g^{+x^k})$$

$$\begin{aligned}
\widehat{R}(g^{+x_k}) &= g(x_k)\widehat{R}(g^{+(x_k,+1)}) + (1-g(x_k))\widehat{R}(g^{+(x_k,-1)}) \\
&= g(x_k)\widehat{R}(g^{+(x_k,+1)}) + (1-g(x_k))\widehat{R}(g) \\
&= g(x_k)\left\{\frac{\pi_p}{n_p}\sum_{x \in X_p \cup \{x_k\}} \{l(g(x),+1) - l(g(x),-1)\} + \frac{1}{n_u}\sum_{x \in X_U} l(g(x),-1)\right\} + (1-g(x_k))\widehat{R}(f)
\end{aligned}$$

constant

Intractable

Solver for PU

$$\widehat{R}_{pu}(g) = \frac{\pi_p}{n_p}\sum_{x\in X_P}[l(g(x),+1) - l(g(x),-1)] + \frac{1}{n_u}\sum_{x\in X_U} l(g(x),-1)$$

$$\ell(t,+1) - \ell(t,-1) = -t$$

如 square loss:

$$l(t,y) = \frac{(t-y)^2}{4} \Rightarrow l(g(x),+1) - l(g(x),-1) = \frac{(g(x)-1)^2 - (g(x)+1)^2}{4} = \frac{-4g(x)}{4} = -g(x)$$

$$\widehat{R}_{pu}(g) = -\frac{\pi_p}{n_p}\sum_{x\in X_P} g(x) + \frac{1}{n_u}\sum_{x\in X_U}\frac{(g(x)+1)^2}{4}$$

$$g(\boldsymbol{x}) = \sum_{j=1}^{b} w_j \phi_j(\boldsymbol{x}) = \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{x}) \qquad \min_{\omega} \widehat{R}(g) + \frac{\lambda}{2}\omega^T\omega$$

$$\min_{\omega} \frac{1}{4n_u}\omega^T\Phi_U^T\Phi_U\omega + \frac{1}{2n_u}1_U^T\Phi_U\omega - \frac{\pi}{n_p}1_P^T\Phi_P\omega + \frac{\lambda}{2}\omega^T\omega$$

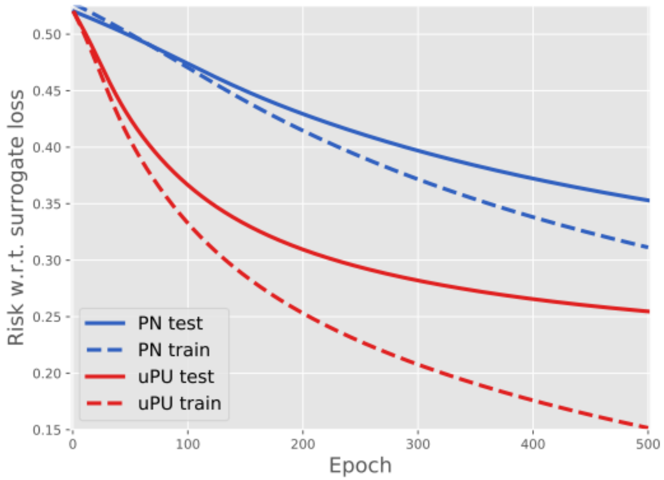
$$\omega = (\frac{1}{2n_u}\Phi_U^T\Phi_U + \lambda I)^{-1}[\frac{\pi}{n_p}\Phi_P^T 1_P - \frac{1}{2n_u}\Phi_U^T 1_U]$$

$\omega : b\times 1$

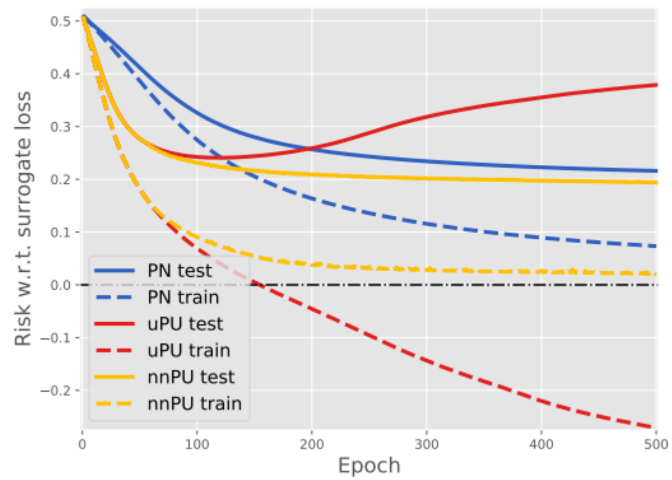
$\Phi_U : u\times b$

$\Phi_P : p\times b$

Over fitting



(a) Plain linear model



(b) Multilayer perceptron (MLP)

$$\pi_{\mathrm{n}}\mathbb{E}_{\mathrm{n}}[\ell(-g(X))] = \underline{\mathbb{E}_X[\ell(-g(X))] - \pi_{\mathrm{p}}\mathbb{E}_{\mathrm{p}}[\ell(-g(X))]}$$

$$\tilde{R}_{\mathrm{pu}}(g) = \frac{\pi_{\mathrm{p}}}{n_{\mathrm{p}}}\sum_{x\in\mathcal{X}_{\mathrm{p}}}[\ell(g(x))] + \max\left\{0, \frac{1}{n_{\mathrm{u}}}\sum_{x\in\mathcal{X}_{\mathrm{u}}}\ell(-g(x)) - \frac{\pi_{\mathrm{p}}}{n_{\mathrm{p}}}\sum_{x\in\mathcal{X}_{\mathrm{p}}}[\ell(-g(x))]\right\}$$