MMD

$$\min_{\boldsymbol{\beta}^T \hat{P}=b} \left\| \frac{1}{p+b}\left(\sum_{x_j \in P}\phi(x_j)+\sum_{x_i \in \hat{P}}\beta_i g_i \phi(x_i)\right)-\frac{1}{\hat{p}-b}\left(\sum_{x_j \in \hat{P}}\phi(x_j)-\sum_{x_i \in \hat{P}}\beta_i g_i \phi(x_i)\right)\right\|_H^2$$

P: 已标记正例集合，$p = |P|$

$\hat{P}$: 预测正例集合，$\hat{p} = |\hat{P}|$

b: 批量查询的个数

$\boldsymbol{\beta}$: $\beta_i \in \{0,1\}$

$\boldsymbol{g}$: $g_i = f(x_i)$

MMD

查询在预测正例中具有代表性的样本

$$\left\| \frac{1}{p+b}\left(\sum_{x_j \in P} \phi(x_j) + \sum_{x_i \in \hat{P}} \beta_i g_i \phi(x_i) \right) - \frac{1}{\hat{p}-b}\left(\sum_{x_i \in \hat{P}} (1-\beta_i g_i)\phi(x_i) \right) \right\|_H^2$$

$$\frac{1}{2}\langle \beta \circ g \rangle^T K_{\hat{P}\hat{P}} \langle \beta \circ g \rangle - \frac{p+b}{p+\hat{p}} 1_{\hat{P}}^T K_{\hat{P}\hat{P}} \langle \beta \circ g \rangle + \frac{\hat{p}-b}{\hat{p}+p} 1_P^T K_{P\hat{P}} \langle \beta \circ g \rangle + const$$

$$\beta_i \in \{0,1\},\ g_i = f(x_i)$$

进一步简写为 $\boldsymbol{\beta}^T K_1 \boldsymbol{\beta} + k\boldsymbol{\beta}$

If Query a Negative Example

现有的PU经验风险公式：

$$R_{emp}(f)=\frac{\pi}{p}\sum_{x_i\in D_P}\tilde{l}(f(x_i),1)+\frac{1}{u}\sum_{x_i\in D_U}l(f(x_i),-1),\ \text{where } p=|D_P|, u=|D_U|$$

$$\Downarrow$$

$$R_{emp}(f)=\frac{\pi}{p}\sum_{x_i\in D_P}\tilde{l}(f(x_i),1)+\frac{1}{u}\sum_{x_i\in D_U}l(f(x_i),-1)+\frac{c}{n}\sum_{x_i\in D_N}l(f(x_i),-1),$$

$$\text{where } p=|D_P|, u=|D_U|, n=|D_N|$$

为了让分类器将已标注的负类样本分对，设置一个较大的c值，加大对负类样本分错的惩罚力度。

$$f^*=\underset{f\in H_k}{arg\,min}\left\{R_{emp}(f)+\lambda\|f\|_{H_k}^2\right\}$$

Cost-sensitive PU Learning

$$\begin{array}{cccc} + & \mathrm{U} & \mathrm{Q} & - \end{array}$$

$$\frac{\pi}{p}\sum_{x_i\in P}\tilde{l}\left(f\left(x_i\right),1\right)+\frac{1}{u}\sum_{x_i\in U\setminus Q}l\left(f\left(x_i\right),-1\right)+\frac{1}{b}\sum_{x_i\in Q}l\left(f\left(x_i\right),-1\right)+\frac{c}{n}\sum_{x_i\in N}l\left(f\left(x_i\right),-1\right)$$

$$=\frac{\pi}{p}\sum_{x_i\in P}\tilde{l}\left(f\left(x_i\right),1\right)+\frac{1}{u}\sum_{x_i\in U}l\left(f\left(x_i\right),-1\right)+\frac{c}{n}\sum_{x_i\in N}l\left(f\left(x_i\right),-1\right)$$

Square Loss

$$l\left(f\left(x\right),y\right)=\frac{1}{4}\left(f\left(x\right)-y\right)^2$$

$$\tilde{l}\left(f\left(x\right),y\right)=l\left(f\left(x\right),y\right)-l\left(f\left(x\right),-y\right)=\frac{1}{4}\left(f\left(x\right)-y\right)^2-\frac{1}{4}\left(f\left(x\right)+y\right)^2=-yf\left(x\right)$$

$$\frac{\pi}{p}\sum_{x_i\in P}\tilde{l}\left(f\left(x_i\right),1\right)+\frac{1}{u}\sum_{x_i\in U}l\left(f\left(x_i\right),-1\right)+\frac{c}{n}\sum_{x_i\in N}l\left(f\left(x_i\right),-1\right)$$

$$\Big\Downarrow$$

$$-\frac{\pi}{p}\sum_{x_i\in P}f\left(x_i\right)+\frac{1}{u}\sum_{x_i\in U}\frac{1}{4}\left(f\left(x_i\right)+1\right)^2+\frac{c}{n}\sum_{x_i\in N}\frac{1}{4}\left(f\left(x_i\right)+1\right)^2$$

Vectorization

$$-\frac{\pi}{p}\sum_{P} f(x_i) + \frac{1}{u}\sum_{U}\frac{1}{4}\left(f(x_i)+1\right)^2 + \frac{c}{n}\sum_{N}\frac{1}{4}\left(f(x_i)+1\right)^2$$

$$\min_{\boldsymbol{\alpha}}\left\{-\frac{\pi}{p}\sum_{x_i\in P}\boldsymbol{\alpha}^T\phi(x_i) + \frac{1}{4u}\sum_{x_i\in U}\left(\boldsymbol{\alpha}^T\phi(x_i)+1\right)^2 + \frac{1}{4n}\sum_{x_i\in N}\left(\boldsymbol{\alpha}^T\phi(x_i)+1\right)^2 + \lambda\|\boldsymbol{\alpha}\|^2\right\}$$

$$\min_{\boldsymbol{\alpha}}\left\{-\frac{\pi}{p}1_P^T\phi_P\boldsymbol{\alpha} + \left(\frac{1}{4u}\boldsymbol{\alpha}^T\phi_U^T\phi_U\boldsymbol{\alpha} + \frac{1}{2u}1_U^T\phi_U\boldsymbol{\alpha}\right) + \left(\frac{c}{4n}\boldsymbol{\alpha}^T\phi_N^T\phi_N\boldsymbol{\alpha} + \frac{c}{2n}1_N^T\phi_N\boldsymbol{\alpha}\right) + \lambda\boldsymbol{\alpha}^T\boldsymbol{\alpha}\right\}$$