References

1. Pattern Classification

   Duda, Hart, Stork

2. Probabilistic theory of Pattern Recognition

   Devroye, Gyorfi, Lugosi

3. Pattern recognition and Machine Learning

   Chris Bishop

$\mathcal{X} \to$ Instance space

$\mathcal{Y} \to$ Label space

Binary Classification

$\mathcal{Y} = \{-1, 1\}$

Classifier

$h : \mathcal{X} \to \{-1, 1\}$ $\quad \mathcal{X} \subseteq \mathbb{R}^d$

Prior

$p_1 = P(Y = 1)$, $p_2 = P(Y = -1)$

Class conditional Distribution

$P(X = x | Y = 1)$, $P(X = x | Y = -1)$

[Densities / p.m.f]

$$\eta(x) = P(Y=1 \mid X=x)$$

$$1-\eta(x) = P(Y=-1 \mid X=x)$$

$$P(Y=1 \mid X=x) = \frac{P(Y=1)\,P(X=x \mid Y=1)}{P(X=x)}$$

Bayes Classifier

$$h(x) = \begin{cases} 1 & P(Y=1 \mid X=x) > P(Y=-1 \mid X=x) \\ -1 & P(Y=-1 \mid X=x) \geq P(Y=1 \mid X=x) \end{cases}$$

$$h(x) = \begin{cases} 1 & \eta(x) > 1-\eta(x) \\ -1 & 1-\eta(x) \geq \eta(x) \end{cases}$$

$$\boxed{h^*(x) = \mathrm{sign}(2\eta(x) - 1)} \rightarrow \text{BC}$$

$$\mathrm{sign}(z) = \begin{cases} 1 & z > 0 \\ -1 & z \leq 0 \end{cases}$$

Specific case

$x \in \mathbb{R}^d$, $\mu \in \mathbb{R}^d$

$C \in \mathbb{R}^{d \times d}$, Symmetric positive semi-definite

$$N(x \mid \mu, C) = \frac{1}{(\sqrt{2\pi})^d |C|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}$$

$|C| \equiv \det(C)$

$$P(X = x \mid Y = 1) = N(x \mid \mu_1, C_1)$$

$$P(X = x \mid Y = -1) = N(x \mid \mu_2, C_2)$$

$$C_1 = C_2 = C$$

$$\eta(x) = \frac{P_1 N(x \mid \mu_1, C_1)}{P(X = x)}$$

$$h^*(x) = \begin{cases} 1 & \log \frac{\eta(x)}{1-\eta(x)} > 0 \\ -1 & \text{otherwise} \end{cases}$$

$$h^*(x) = \operatorname{sign}\left(\log \frac{\eta(x)}{1-\eta(x)}\right)$$

$$\log \frac{\eta(x)}{1-\eta(x)} = \log \frac{P_1}{P_2} \frac{N(x \mid \mu_1, C)}{N(x \mid \mu_2, C)}$$

$$= \log \frac{P_1}{P_2} + \log \frac{N(x \mid \mu_1, C)}{N(x \mid \mu_2, C)}$$

$$\log \frac{N(x \mid \mu_1, C)}{N(x \mid \mu_2, C)}$$

$$= -\frac{1}{2}(x-\mu_1)^T C^{-1}(x-\mu_1)$$

$$+\frac{1}{2}(x-\mu_2)^T C^{-1}(x-\mu_2)$$

$$= (\mu_1-\mu_2)^T C^{-1} x$$

$$-\frac{1}{2}\left(\mu_1^T C^{-1}\mu_1 - \mu_2^T C^{-1}\mu_2\right)$$

$$\log \frac{\eta(x)}{1-\eta(x)}$$

$$= (\mu_1 - \mu_2)^T C^{-1} x$$

$$- \frac{1}{2}\left(\mu_1^T C^{-1} \mu_1 - \mu_2^T C^{-1} \mu_2\right) + \log \frac{p_1}{p_2}$$

$$= w^T x + b$$

$$w = C^{-1}(\mu_1 - \mu_2)$$

$$b = \log \frac{p_1}{p_2} - \frac{1}{2}\left(\mu_1^T C^{-1} \mu_1 - \mu_2^T C^{-1} \mu_2\right)$$

$$h(x) = \operatorname{sign}(w^T x - b)$$

Bayes Classifier is a Linear function

Examine the case

$C_1 \neq C_2$, $C_1 = \sigma_1^2 I$
$C_2 = \sigma_2^2 I$

$w^T x = b$ hyperplane

$w^T x > b$ halfspace

$w^T x < b$ halfspace

How good is the Bayes Classifier (BC)

$$E(\mathbb{1}_A) = P(A)$$

$$P(h(X) \neq Y)$$

$$= E_{XY} \mathbb{1}\{h(X) \neq Y\}$$

$$= E_X E_{Y|X} \mathbb{1}\{h(X) \neq Y\}$$

$$E_{Y|\cancel{X}} \mathbb{1}\{h(x) \neq Y\}$$

$$= \eta(x) \mathbb{1}\{h(x) \neq 1\} + (1-\eta(x))\mathbb{1}\{h(x) \neq -1\}$$

$$P(h(X) \neq Y) =$$

$$E_X\left(\eta(X)\,\mathbb{1}_{\{h(X) \neq 1\}} + (1-\eta(X))\,\mathbb{1}_{\{h(X) \neq -1\}}\right)$$

$$P(h(X) \neq Y) - P(h^*(X) \neq Y)$$

$$= E_X\left(E_{Y|X}\left(\mathbb{1}_{\{h(X) \neq Y\}} - \mathbb{1}_{\{h^*(X) \neq Y\}}\right)\right)$$

$$h(X) \neq h^*(X)$$

$$\Rightarrow h(X) \neq Y \Rightarrow h^*(X) = Y$$

$$E_{Y|X}\left(\mathbb{1}_{\{h(X)\neq Y\}} - \mathbb{1}_{\{h^*(X)\neq Y\}}\right)$$

$$= E_{Y|X}\left(\mathbb{1}_{\{h^*(X)=Y\}} - \mathbb{1}_{\{h^*(X)\neq Y\}}\right)$$

$$= E_{Y|X}\left(2\mathbb{1}_{\{h^*(X)=Y\}} - 1\right)$$

$$= \eta(x)\left(2\mathbb{1}_{\{h^*(x)=1\}} - 1\right)$$

$$+ (1-\eta(x))\left(2\mathbb{1}_{\{h^*(x)=-1\}} - 1\right)$$

$$= \eta(x)\left(2\mathbb{1}_{\{h^*(x)=1\}} - 1\right)$$

$$+ (1-\eta(x))\left(2\left(1 - \mathbb{1}_{\{h^*(x)=1\}}\right)\right)$$

-

$$= (2\eta(x)-1)\left(2\mathbb{1}_{\{h^*(x)=1\}} - 1\right)$$

$\eta(x)$	$(2\eta(x)-1)$	$2\mathbb{1}_{\{h^*(x)=1\}} - 1$
$> \frac{1}{2}$	> 0	≥ 0
$\leq \frac{1}{2}$	≤ 0	< 0

$$E_{Y|X}\left(\mathbb{1}\{h(X) \neq Y\} - \mathbb{1}\{h^*(X) \neq Y\}\right)$$

$$= (2\eta(x)-1)\left(2\mathbb{1}\{h^*(x)=1\} - 1\right)$$

$$= |2\eta(x) - 1|$$

$$P(h(X) \neq Y) - P(h^*(X) \neq Y)$$

$$= E_X |2\eta(X) - 1| \geq 0$$

[Thm 2-1 DGL]