9/22/19
CSE176

CSE176 Homework #1

#1 $P[D] = 0.05$ => Athletes use drugs

$P[D'] = 1 - 0.05 = 0.95$ => Athletes don't use drugs

$P[FP] = 0.02$ => false positive

$P[FP'] = 0.98$ => True positive

$P[FN] = 0.15$ => false Negative

$P[FN'] = 0.85$ => True Negative

1.) $P[D|FP'] = \frac{P[FP'|D]\,P[D]}{P[FP']} = \frac{((0.98)(0.05)) \cdot (0.05)}{(0.98} = 0.0025$ or $\boxed{.25\%}$

2.) $P[D'|FN'] = \frac{P[FN'|D']\,P[D']}{P[FN']} = \frac{((0.85)(0.95))\,(0.95)}{(0.85)} = 0.9025$ or $\boxed{90.25\%}$

ise #2

1.) $R_i(x) = \sum_{K=1}^{K} \lambda_{iK}\, P[C_K|x]$

Choosing class : $\arg\min_{K=1,\dots,K} R_K(x)$

2.) $\operatorname{argmax}\{P(C_1|x), \dots, P(C_K|x), 1 - \lambda_{21}\}$

$\lambda_{21} P(C_1|x) > P(C_2|x)$

3.) $R_1(x) = 1 - P(C_1|x)$

4.) $\lambda_{21} = 0.99$ b/c This will makes us choose class 1 always unless $P(C_2|x) > .99$

ercise #3

Association Rule	Support	Confidence
meat → Avocado	$3/6 = 1/2$	$3/5$
avocado → MEAT	$3/6 = 1/2$	$3/4$
yogurt → Avocado	$2/6 = 1/3$	$2/3$
avocado → yogurt	$2/6 = 1/3$	$2/4 = \frac{1}{2}$
meat → yogurt	$2/6 = 1/2$	$2/5$
yogurt → MEAT	$2/6 = 1/2$	$2/3$

The best rule to use is

Exercise #4

$\theta = 0.6$

1.

	predicted NO	predicted yes	
Actual NO	1	2	3
Actual yes	0	2	2
	1	4	

$\frac{(FP+FN)}{total} = \frac{0+2}{5} = \frac{2}{5} =$

Classification error : 40%

$\theta = 0.7$

	predicted NO	predicted yes	
Actual NO	1	2	3
Actual yes	1	1	2
	2	3	

$\frac{1+2}{5} = \frac{3}{5}$

Classification error : 60 %

$\theta = 0.5$

	predicted NO	predicted yes	
Actual NO	2	1	3
Actual yes	0	2	2
	2	3	

$\frac{0+1}{5} = \frac{1}{5}$

Classification error : 20 %

$\theta = 0.9$

	predicted NO	predicted yes	
Actual NO	0	3	3
Actual yes	1	1	2
	1	4	

$\frac{3+1}{5} = \frac{4}{5}$

Classification error : 80 %

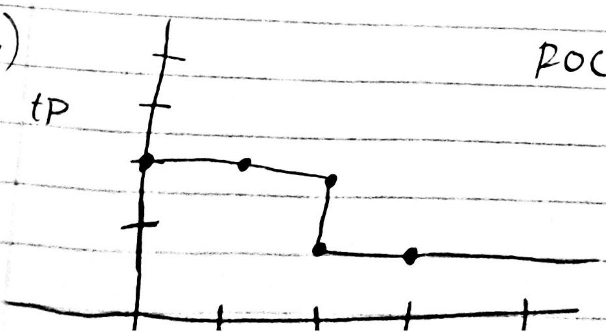
$\theta = 0.2$

	predicted NO	predicted yes	
Actual NO	3	0	3
Actual yes	0	2	2
	3	2	

$\frac{0}{5} = 0\%$ Classification error

2.)

ROC Curve



Exercise #5

Classifier A has $fp_A, tp_A \in [0,1]$ $fp_A > tp_A$

Classifier B has $fp_B, tp_B \in [0,1]$ $tp_B > fp_B$

The point is located above the diagonal on the ROC space.

$$\text{tp-rate} = \frac{tp_B}{p} \qquad \text{fp-rate} = \frac{tp_B}{n}$$

#6

$$\Theta = \{\theta_1, \theta_2, \theta_3\} \subset \mathbb{R} \qquad x \in \mathbb{R}$$

$$h(x;\Theta) = \theta_1 + \theta_2 \sin 2x + \theta_3 \sin 4x \in \mathbb{R}$$

1.) $E(h;\Theta) = \frac{1}{N}\sum_{n=1}^{N} (y_n - h(x_n))^2$

2.) $E(h;\Theta) = \left(\frac{1}{N}\sum_{n=1}^{N} (y_n - \theta_1 + \theta_2 \sin 2x_n + \theta_3 \sin 4x_n)^2\right) \frac{\partial \Theta}{\partial x}$

$$\frac{2}{N}\sum_{n=1}^{N} (y_n - \theta_1 + \theta_2 \sin 2x_n + \theta_3 \sin 4x_n)\cdot(2\cos(2x))\cdot(4\cos(4x))$$

$$\frac{16}{N}\cos(2x_n)\cos(4x_n)\sum_{n=1}^{N} (y_n - \theta_1 + \theta_2 \sin 2x_n + \theta_3 \sin 4x_n)$$

3.) $h(x;\theta_1,\theta_2,\theta_3) = \theta_1 x^2 + \theta_2 x + \theta_3$

$\Theta = A^{-1}y$, $A = \begin{vmatrix} 1 & 1 & \frac{1}{N}\sum_{n=1}^{N} x_n \\ 1 & \frac{1}{N}\sum_{n=1}^{N} x_n & \frac{1}{N}\sum_{n=1}^{N} x_n^2 \\ \frac{1}{N}\sum_{n=1}^{N} x_n & \frac{1}{N}\sum_{n=1}^{N} x_n^2 & \frac{1}{N}\sum_{n=1}^{N} x_n^3 \end{vmatrix}$

$\Theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$

inverse

$y = \begin{vmatrix} \frac{1}{N}\sum_{n=1}^{N} y_n \\ \frac{1}{N}\sum_{n=1}^{N} y_n x_n \\ \frac{1}{N}\sum_{n=1}^{N} y_n x_n^2 \end{vmatrix}$

$\Theta = A^{-1}y$ multiply y

4.) $\int_0^{2\pi} \frac{1}{N} \sum_{n=1}^{N} f(x_n)\, dx = \frac{1}{N} \sum_{n=1}^{N} \int_0^{2\pi} f(x_n)\, dx$

Exercise #7

$x \in \{0, 1, 2, 3, \cdots\}$ Poisson Distribution w/ mass fun

$$p(x; \theta) = \frac{e^{-\theta} \theta^{x}}{x!}$$

1.) $\sum_{x=0}^{\infty} p(x) = \int_0^{\infty} p'(x)\, dx \rightarrow p'(x) = \lim_{h \to 0} \frac{p(x+h) - p(x)}{h}$

$= \frac{0}{0} =$ l'hop $= \lim_{h \to 0} p'(x+h) - p'(x) = 0$

2.) $\mathcal{L}(\Theta; X) = \log \prod_{n=1}^{N} p(x_n; \Theta) = \sum_{n=1}^{N} \log p(x_n; \Theta)$

3.) $\sum_{n=1}^{N} \log\left(\frac{e^{-\theta} \theta^{x}}{x!}\right) = \sum_{n=1}^{N} \log(e^{-\theta} \theta^{x}) - \log(x!)$

$= \sum_{n=1}^{N} \log(e^{-\theta}) + \log(\theta^{x}) - \log(x!)$

$= \sum_{n=1}^{N} -\theta \log(e) + x \log(\theta) - \log(x!)$

4.) Maximum log likelihood $\Theta_{MLE} = \arg\max_{\Theta} \mathcal{L}(\Theta; X)$

poisson

$\Theta = \{\theta\}$, $p(x; \theta)$.

$\mathcal{L}\{\theta\}$ $\frac{1}{N} \sum_{n=1}^{N} x_n$

EXERCISE #8

$\theta \in [0,1]^D$ $\quad x \in \{0,1\}^D$ $\quad$ Binary Random vector i

$$p(x;\theta) = \prod_{d=1}^{D} \theta_d^{x_d} (1-\theta_d)^{1-x_d}$$

1.) $p(x;\theta) = \theta^x (1-\theta)^{1-x}$

$\max_\theta \mathcal{L}(\theta) = \sum_{n=1}^{N} \log(\theta^x \cdot (1-\theta)^{1-x}) =$

$\sum_{n=1}^{N} \log(\theta^x) + \log((1-\theta)^{1-x}) =$

$\sum_{n=1}^{N} x\log(\theta) + (1-x)\log(1-\theta) =$

Take partial derivative of θ $\quad \frac{\partial \mathcal{L}}{\partial \theta}$

$\frac{\sum x_n}{\theta} - \frac{n - \sum x_n}{1-\theta}\ 0 \rightsquigarrow \frac{(1-\theta)\sum x_n - \theta(n - \sum x_n)}{\theta(1-\theta)} = 0$

$\rightsquigarrow (1-\theta)\sum x_n - \theta(n-\sum x_n) = \sum x_n - \theta\sum x_n - n\theta + \theta\sum x_n = 0$

$\rightsquigarrow \sum x_n - \theta\sum x_n = n\theta - \theta\sum x_n$

$\hat{\theta} = \frac{1}{n}\sum x_n$

2.) $g_k(x) = p(x|C_k)\, p(C_k)$ $\quad$ or $\quad$; $g_k(x) = w_k^T x + w_{k0}$ Bernoulli

or $\quad g_k(x) = \log p(x|C_k) + \log p(C_k)$ $\quad$ rule $w^T x + w_0 > 0$

3.) $p(x|C_k) \cdot \pi_k = g_k(x)$

$$\pi_k^a \prod_{d=1}^{D} \theta_d^{x_d} (1-\theta_d)^{1-x_d}$$

4.) $w^T x + w_0 > 0$

5.)

EXERCISE # 9

$x \in \mathbb{R}^D$ $\quad p(x \mid C_1) \sim \mathcal{N}(\mu, \sigma_1^2 I)$ $\quad p(x \mid C_2) \sim \mathcal{N}(\mu, \sigma_2^2 I)$

Class boundary =

$$\frac{\text{upper class limit of lower class} - \text{lower class limit of upper class}}{2}$$

it has a bell shape Curve