Pedro Nunes 109368
Andrei Barb 109762

1) a.

We want to calculate the necessary parameters for

class $\leftarrow$ $\rightarrow$ observation

$$P(c|x) = \frac{P(x|c) \times P(c)}{P(x)}$$

, where $P(x)$ is optional for classification purposes.

$P(c)$: $P(N) = 3/6 = 1/2$ $P(P) = 3/6 = 1/2$

$P(y_3, y_4 | C)$: $P(0,0|N) = 1/3$ $P(0,0|P) = 2/3$
$P(0,1|N) = 2/3$ $P(0,1|P) = 0$
$P(1,0|N) = 0$ $P(1,0|P) = 1/3$
$P(1,1|N) = 0$ $P(1,1|P) = 0$

$P(y_5 | C)$: $P(0|N) = 1/3$ $P(0|P) = 0$
$P(1|N) = 2/3$ $P(1|P) = 3/3 = 1$

$P(y_1, y_2 | C) \sim N(\mu, \Sigma | C)$

$$\Sigma = \begin{bmatrix} var(y_1) & cov(y_2, y_1) \\ cov(y_1, y_2) & var(y_2) \end{bmatrix}$$

$$\mu = \begin{bmatrix} E(y_1) \\ E(y_2) \end{bmatrix}$$

$P(y_1, y_2 | N)$:

$$E(y_1) = \frac{0,52 + 0,53 + 0,42}{3} = 0,49$$

$$E(y_2) = \frac{0,80 + 0,92 + 0,48}{3} = 0,7(3) \simeq 0,73$$

$$\mathrm{var}(y_1) = \frac{\sum (y_{1i} - \mu)^2}{m-1} = \frac{(0,52-0,49)^2 + (0,53-0,49)^2 + (0,42-0,49)^2}{3-1}$$

$$= 0,0037$$

$$\mathrm{var}(y_2) = \frac{(0,8-0,73)^2 + (0,92-0,73)^2 + (0,48-0,73)^2}{3-1} = 0,05175$$

$$\mathrm{cov}(y_1, y_2) = \frac{\sum (y_{1i} - \mu_1)(y_{2i} - \mu_2)}{m-1}$$

$$= \frac{(0,52-0,49)(0,8-0,73) + (0,53-0,49)(0,92-0,73)}{3-1} +$$

$$+ \frac{(0,42-0,49)(0,48-0,73)}{3-1} = 0,0136$$

$$P(y_1, y_2 | N) \sim N\left(\mu = \begin{bmatrix} 0,49 \\ 0,73 \end{bmatrix}, \Sigma = \begin{bmatrix} 0,0037 & 0,0136 \\ 0,0136 & 0,05175 \end{bmatrix} \middle| N\right)$$

$$\det(\Sigma | N) = 0,0037 \times 0,05175 - 0,0136 \times 0,0136 = 6,515 \times 10^{-6}$$

$$(\Sigma^{-1} | N) = \frac{1}{6,515 \times 10^{-6}} \begin{bmatrix} 0,05175 & -0,0136 \\ -0,0136 & 0,0037 \end{bmatrix} \simeq \begin{bmatrix} 7943,2 & -2087,5 \\ -2087,5 & 567,9 \end{bmatrix}$$

$P(y_1, y_2 \mid P)$:

$$E(y_1) = \frac{0{,}49 + 0{,}62 + 0{,}44}{3} = 0{,}51(6) \simeq 0{,}5167$$

$$E(y_2) = \frac{0{,}58 + 0{,}31 + 0{,}38}{3} = 0{,}42(3) \simeq 0{,}4233$$

$$\mathrm{var}(y_1) = \frac{(0{,}49-0{,}5167)^2 + (0{,}62-0{,}5167)^2 + (0{,}44-0{,}5167)^2}{2}$$

$$= 0{,}00863$$

$$\mathrm{Var}(y_2) = \frac{(0{,}58-0{,}4233)^2 + (0{,}31-0{,}4233)^2 + (0{,}38-0{,}4233)^2}{2}$$

$$= 0{,}019633335$$

$$\mathrm{cov}(y_1, y_2) = \frac{(0{,}49-0{,}5167)(0{,}58-0{,}4233) + (0{,}62-0{,}5167)(0{,}31-0{,}4233)}{2}$$

$$+ \frac{(0{,}44-0{,}5167)(0{,}38-0{,}4233)}{2} = -0{,}00628$$

$$P(y_1, y_2 \mid P) \sim N\left(\mu = \begin{bmatrix} 0{,}5167 \\ 0{,}4233 \end{bmatrix}, \Sigma = \begin{bmatrix} 0{,}00863 & -0{,}00628 \\ -0{,}00628 & 0{,}019633335 \end{bmatrix} \mid P \right)$$

$$\det(\Sigma \mid P) = 0{,}00863 \times 0{,}019633335 - (-0{,}00628 \times (-0{,}00628)) \simeq 1{,}3 \times 10^{-4}$$

$$(\Sigma^{-1} \mid P) = \frac{1}{1{,}3 \times 10^{-4}} \begin{bmatrix} 0{,}019633335 & 0{,}00628 \\ 0{,}00628 & 0{,}00863 \end{bmatrix} \simeq \begin{bmatrix} 151 & 48{,}3 \\ 48{,}3 & 66{,}38 \end{bmatrix}$$

b.

$$x_7 : P(x_7 | N) = P(0{,}45 ; 0{,}80 ; 0 ; 0 ; 1 | N)$$

$$= P(0{,}45 ; 0{,}80 | N)\, P(0{,}0 | N)\, P(1 | N)$$

$$(x-\mu)^T = \begin{bmatrix} 0{,}45 - 0{,}49 \\ 0{,}80 - 0{,}73 \end{bmatrix}^T = \begin{bmatrix} -0{,}04 & 0{,}07 \end{bmatrix}$$

$P(0{,}45 , 0{,}8 | N) =$

$$= \frac{1}{(2\pi)^{2/2} \times \sqrt{6{,}515 \times 10^{-6}}} \times e^{-\frac{1}{2}\left(\begin{bmatrix} -0{,}04 & 0{,}07 \end{bmatrix} \begin{bmatrix} 7943{,}2 & -2087{,}5 \\ -2087{,}5 & 567{,}9 \end{bmatrix} \begin{bmatrix} -0{,}04 \\ 0{,}07 \end{bmatrix} \right)}$$

$$= 7{,}81 \times 10^{-5}$$

$$P(x_7 | N) = 7{,}81 \times 10^{-5} \times \frac{1}{3} \times \frac{2}{3} = 1{,}74 \times 10^{-5}$$

$$P(N | x_7) = \frac{P(x_7 | N) \times P(N)}{P(x_7)} = \frac{1{,}74 \times 10^{-5} \times 1/2}{P(x_7)} = \frac{8{,}68 \times 10^{-6}}{P(x_7)}$$

$$P(x_7 | P) = P(0{,}45;\ 0{,}80;\ 0;\ 0;\ 1 | P)$$

$$= P(0{,}45;\ 0{,}80 | P)\ P(0{,}0 | P)\ P(1 | P)$$

$$(x-\mu)^T = \begin{bmatrix} 0{,}45 - 0{,}5167 \\ 0{,}80 - 0{,}4233 \end{bmatrix}^T = \begin{bmatrix} -0{,}0667 \\ 0{,}3767 \end{bmatrix}^T$$

$P(0{,}45,\ 0{,}80 | P) =$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{1{,}3\times 10^{-4}}} \cdot e^{-\frac{1}{2}\left(\begin{bmatrix} -0{,}0667 & 0{,}3767 \end{bmatrix}' \begin{bmatrix} 151 & 48{,}3 \\ 48{,}3 & 66{,}38 \end{bmatrix} \begin{bmatrix} -0{,}0667 \\ 0{,}3767 \end{bmatrix}\right)}$$

$\simeq 0{,}30.$

$$P(x_7 | P) = 0{,}30 \times \frac{2}{3} \times 1 = 0{,}2$$

$$P(P | x_7) = \frac{P(x_7 | P) \times P(P)}{P(x_7)} = \frac{0{,}2 \times 1/2}{P(x_7)} = \frac{0{,}1}{P(x_7)}$$

$$\boxed{P(N | x_7) < P(P | x_7) \iff \frac{8{,}68\times 10^{-6}}{\cancel{P(x_7)}} < \frac{0{,}1}{\cancel{P(x_7)}} \iff 8{,}6\times 10^{-6} < 0{,}1}$$

Therefore, we classify x_7 as P

$x_8: \; P(x_8 | N) = P(0{,}50; 0{,}70; 0; 1; 1 | N)$

$$= P(0{,}50; 0{,}70 | N)\, P(0; 1 | N)\, P(1 | N)$$

$$(x-\mu)^T = \begin{bmatrix} 0{,}50 - 0{,}49 \\ 0{,}70 - 0{,}73 \end{bmatrix}^T = \begin{bmatrix} 0{,}01 \\ -0{,}03 \end{bmatrix}^T$$

$P(0{,}50; 0{,}70 | N) =$

$$= \frac{1}{(2\pi)^{2/2}\sqrt{6{,}515 \times 10^{-6}}} \cdot e^{-\frac{1}{2}\left(\begin{bmatrix} 0{,}01 & -0{,}03 \end{bmatrix} \begin{bmatrix} 7943{,}2 & -2087{,}5 \\ -2087{,}5 & 567{,}9 \end{bmatrix} \begin{bmatrix} 0{,}01 \\ -0{,}03 \end{bmatrix}\right)}$$

$\simeq 17{,}35$

$P(x_8 | N) = 17{,}35 \times 2/3 \times 2/3 = 7{,}7(1) \simeq 7{,}7$

$$P(N | x_8) = \frac{P(x_8 | N)\, P(N)}{P(x_8)} = \frac{7{,}7 \times 1/2}{P(x_8)} = \frac{3{,}85}{P(x_8)}$$

$$P(x_8 \mid P) = P(0{,}50;\, 0{,}70;\, 0;\, 1;\, 1 \mid P)$$

$$= P(0{,}50;\, 0{,}70 \mid P)\, P(0;\, 1 \mid P)\, P(1 \mid P)$$

$$(x-\mu)^T = \begin{bmatrix} 0{,}50 - 0{,}5167 \\ 0{,}70 - 0{,}4233 \end{bmatrix}^T = \begin{bmatrix} -0{,}0167 \\ 0{,}2767 \end{bmatrix}^T$$

$$P(0{,}50;\, 0{,}70 \mid P) =$$

$$= \frac{1}{(2\pi)^{1/2} \times \sqrt{1{,}3 \times 10^{-4}}} \cdot e^{-\frac{1}{2}\left(\begin{bmatrix} -0{,}0167 & 0{,}2767 \end{bmatrix} \begin{bmatrix} 151 & 48{,}3 \\ 48{,}3 & 66{,}38 \end{bmatrix} \begin{bmatrix} -0{,}0167 \\ 0{,}2767 \end{bmatrix}\right)}$$

$$= 1{,}346$$

$$P(x_8 \mid P) = 1{,}346 \times 0 \times 1 = 0$$

$$P(P \mid x_8) = \frac{P(x_8 \mid P)\, P(P)}{P(x_8)} = \frac{0}{P(x_8)}$$

$$\boxed{P(N \mid x_8) > P(P \mid x_8) \iff \frac{3{,}85}{\cancel{P(x_8)}} > \frac{0}{\cancel{P(x_8)}} \iff 3{,}85 > 0}$$

Therefore, we classify x_8 as N

2)

$x_1 : (0,1,1) \rightarrow z = N$

hamming (x_1, x_2) = hamming $((0,1,1), (0,0,0)) = 2$
hamming $(x_1, x_3) = 0$
hamming $(x_1, x_4) = 2$
hamming $(x_1, x_5) = 1$
hamming $(x_1, x_6) = 1$
hamming $(x_1, x_7) = 1$
hamming $(x_1, x_8) = 0$

There is no radius with only $K=3$ so we use $K=2$.

$K=2 \rightarrow (x_3, x_8) \Rightarrow \hat{z} = moda(N,N) = N$
$\hat{z} = z$ ✓

$x_2 : (0,0,0) \rightarrow z = N$

hamming $(x_2\ x_1) = 2$
hamming $(x_2, x_3) = 2$
hamming $(x_2, x_4) = 2$
hamming $(x_2, x_5) = 1$
hamming $(x_2, x_6) = 1$
hamming $(x_2, x_7) = 1$
hamming $(x_2, x_8) = 2$

$K=3 \rightarrow (x_5, x_6, x_7) \Rightarrow \hat{z} = moda(P,P,P) = P$
$z \neq \hat{z}$ ✗

$x_3 : (0,1,1) \rightarrow z = N$

hamming $(x_3, x_1) = 0$
hamming $(x_3, x_2) = 2$
hamming $(x_3, x_4) = 2$
hamming $(x_3, x_5) = 1$
hamming $(x_3, x_6) = 1$
hamming $(x_3, x_7) = 1$
hamming $(x_3, x_8) = 0$

There is no radius with only $K=3$ so we use $K=2$.

$K=2 \rightarrow (x_1, x_8) \Rightarrow \hat{z} = \text{moda}(N,N) = N$
$\hat{z} = z$ ✓

$x_4 : (1,0,1) \rightarrow z = P$

hamming $(x_4, x_1) = 2$
hamming $(x_4, x_2) = 2$
hamming $(x_4, x_4) = 2$
hamming $(x_4, x_5) = 1$
hamming $(x_4, x_6) = 1$
hamming $(x_4, x_7) = 1$
hamming $(x_4, x_8) = 2$

$K=3 \rightarrow (x_5, x_6, x_7) \Rightarrow \hat{z} = \text{moda}(P,P,P) = P$
$\hat{z} = z$ ✓

$x_5 : (0,0,1) \rightarrow z = P$

hamming $(x_5, x_1) = 1$
hamming $(x_5, x_2) = 1$
hamming $(x_5, x_3) = 1$
hamming $(x_5, x_4) = 1$
hamming $(x_5, x_6) = 0$
hamming $(x_5, x_7) = 0$
hamming $(x_5, x_8) = 1$

There is no radius with only $k=3$, so we use $k=2$

$k=2 \rightarrow (x_6, x_7) \Rightarrow \hat{z} = \text{moda}(P, P) = P$
$\hat{z} = z$ ✓

$x_6 : (0,0,1) \rightarrow z = P$

hamming $(x_6, x_1) = 1$
hamming $(x_6, x_2) = 1$
hamming $(x_6, x_3) = 1$
hamming $(x_6, x_4) = 1$
hamming $(x_6, x_5) = 0$
hamming $(x_6, x_7) = 0$
hamming $(x_6, x_8) = 1$

There is no radius with only $k=3$, so we use $k=2$

$k=2 \rightarrow (x_5, x_7) \Rightarrow \hat{z} = \text{moda}(P, P) = P$
$\hat{z} = z$ ✓

$x_7 : (0,0,1) \longrightarrow z = P$

$$\text{hamming}(x_7, x_1) = 1$$
$$\text{hamming}(x_7, x_2) = 1$$
$$\text{hamming}(x_7, x_3) = 1$$
$$\text{hamming}(x_7, x_4) = 1$$
$$\text{hamming}(x_7, x_5) = 0$$
$$\text{hamming}(x_7, x_6) = 0$$
$$\text{hamming}(x_7, x_8) = 1$$

There is no radius with only $k=3$, so we use $k=2$

$k=2 \longrightarrow (x_5, x_6) \Rightarrow \hat{z} = \text{moda}(P,P) = P$
$\hat{z} = z$ ✓

$x_8 : (0,1,1) \longrightarrow z = N$

$$\text{hamming}(x_8, x_1) = 0$$
$$\text{hamming}(x_8, x_2) = 2$$
$$\text{hamming}(x_8, x_3) = 0$$
$$\text{hamming}(x_8, x_4) = 2$$
$$\text{hamming}(x_8, x_5) = 1$$
$$\text{hamming}(x_8, x_6) = 1$$
$$\text{hamming}(x_8, x_7) = 1$$

There is no radius with only $k=3$, so we use $k=2$

$k=2 \longrightarrow (x_1, x_3) \Rightarrow \hat{z} = \text{moda}(N,N) = N$
$\hat{z} = z$ ✓

$$\boxed{\text{Accuracy} = \frac{7}{8} = 0{,}875 = 87{,}5\%}$$

3) a. $\bullet P(\theta=0)=p \Rightarrow P(\theta=1)=1-p$

$\bullet P(X=x \mid \theta=0) = P(X=x \mid \theta=1)$

Teorema de Bayes

$$P(\theta=0 \mid X=x) = \frac{P(X=x \mid \theta=0)\,P(\theta=0)}{P(X=x)}$$

$$\Leftrightarrow P(\theta=0 \mid X=x) = \frac{p\,P(X=x \mid \theta=0)}{P(X=x)}$$

$$\Leftrightarrow P(X=x \mid \theta=0) = \frac{P(X=x)\,P(\theta=0 \mid X=x)}{p}$$

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

$$\Leftrightarrow P(\theta=1 \mid X=x) = \frac{(1-p)\,P(X=x \mid \theta=1)}{P(X=x)}$$

$$\Leftrightarrow P(X=x \mid \theta=1) = \frac{P(X=x)\,P(\theta=1 \mid X=x)}{(1-p)}$$

$$P(X=x \mid \theta=0) = P(X=x \mid \theta=1)$$

$$\Leftrightarrow \frac{P(X=x)\,P(\theta=0 \mid X=x)}{p} = \frac{P(X=x)\,P(\theta=1 \mid X=x)}{1-p}$$

$$\Leftrightarrow \frac{P(\theta=0 \mid X=x)}{P(\theta=1 \mid X=x)} = \frac{p}{1-p}$$

Because $p \in (1/2, 1]$, we know that $P(\theta=0 \mid X=x) > P(\theta=1 \mid X)$

Therefore the MAP classifier always predicts $\theta_{Bayes}=0$

$$E_{Bayes} = P(\theta \neq \theta_{Bayes}) = P(\theta \neq 0) = P(\theta=1) = 1-p$$

b. we know from a. that:

- $P(\theta=0 \mid X=x) = \dfrac{P(X=x \mid \theta=0)\, P(\theta=0)}{P(X=x)}$

- $P(\theta=1 \mid X=x) = \dfrac{P(X=x \mid \theta=1)\, P(\theta=1)}{P(X=x)}$

- $P(X=x \mid \theta=0) = P(X=x \mid \theta=1) = k$

- $P(\theta=0) = p \Rightarrow P(\theta=1) = 1-p$

Lei probabilidade Total

$$P(X=x) = P(X=x \mid \theta=0)\, P(\theta=0) + P(X=x \mid \theta=1)\, P(\theta=1)$$
$$= kp + k(1-p)$$

$$2 \times P(\theta=0 \mid X=x) \times P(\theta=1 \mid X=x) =$$

$$= 2 \times \frac{P(X=x \mid \theta=0)\, P(\theta=0)}{P(X=x)} \times \frac{P(X=x \mid \theta=1)\, P(\theta=1)}{P(X=x)}$$

$$= 2 \times \frac{kp}{kp+k(1-p)} \times \frac{k(1-p)}{kp+k(1-p)}$$

$$= 2 \times \frac{k^2 p(1-p)}{(kp+k(1-p))^2} = 2 \times \frac{k^2 p(1-p)}{(k(p+(1-p)))^2} = 2 \times \frac{\cancel{k^2}\, p(1-p)}{\cancel{k^2}\,(p+(1-p))^2}$$

$$= 2 \times \frac{p(1-p)}{(p+1-p)^2} = 2 \times \frac{p(1-p)}{1^2} = 2p(1-p)$$

$$E_{1NN} = 2p(1-p)$$

c. From before:

- $E_{1NN} = 2p(1-p)$
- $E_{Bayes} = 1-p$

$E_{Bayes} = 1-p \Leftrightarrow p = 1 - E_{Bayes}$

$$E_{1NN} = 2p(1-p) = 2(1-E_{Bayes})(1-(1-E_{Bayes}))$$
$$= 2(1-1+E_{Bayes})(1-E_{Bayes})$$
$$= 2E_{Bayes}(1-E_{Bayes})$$