Aprendizagem 2021/22

Homework I – Group 013

I. Pen-and-paper

1)

Priors :

- *Class 0*: $P(Class=0)=\frac{\#class=0}{\#total}=\frac{4}{10}$
- *Class 1*: $P(Class=1)=\frac{\#class=1}{\#total}=\frac{6}{10}$

Y1 distribution:

- *Class 0*: $\mu=\frac{1}{4}\sum_{i=1}^{4}x_i=0.25 \quad \sigma=\sqrt{\frac{1}{4-1}\sum_{i=1}^{4}(x_i-\mu)^2}=0.2380 \Rightarrow$
  $(y1|class=0)\sim N(\mu,\sigma^2)\sim N(0.25,0.2380^2)$
- *Class 1*: $\mu=\frac{1}{6}\sum_{i=5}^{10}x_i=0.05 \quad \sigma=\sqrt{\frac{1}{6-1}\sum_{i=5}^{10}(x_i-\mu)^2}=0.2881 \Rightarrow$
  $(y1|class=1)\sim N(\mu,\sigma^2)\sim N(0.05,0.2881^2)$

Y2 probability mass function:

- *Class 0*: $P(y2\mid Class=0)=\begin{cases}\frac{\#(y2=A\mid class=0)}{\#class=0}=\frac{2}{4}, & if\ y2=A\\ \frac{\#(y2=B|class=0)}{\#class=0}=\frac{1}{4}, & if\ y2=B\\ \frac{\#(y2=C|class=0)}{\#class=0}=\frac{1}{4}, & if\ y2=C\end{cases}$

-

- *Class 1*: $P(y2\mid Class=1)=\begin{cases}\frac{\#(y2=A\mid class=1)}{\#class=1}=\frac{1}{6}, & if\ y2=A\\ \frac{\#(y2=B|class=1)}{\#class=1}=\frac{2}{6}, & if\ y2=B\\ \frac{\#(y2=C|class=1)}{\#class=1}=\frac{3}{6}, & if\ y2=C\end{cases}$

Y3 and Y4 distribution:

- *Class 0*: $\mu=\frac{1}{4}\sum_{i=1}^{4}[y_{3_i}\quad y_{4_i}]=[0.2\quad 0.25]$ $\Sigma=\begin{bmatrix}cov(y3,y3) & cov(y3,y4)\\ cov(y4,y3) & cov(y4,y4)\end{bmatrix}=\begin{bmatrix}0.18 & 0.18\\ 0.18 & 0.25\end{bmatrix}$

  $det(\Sigma)=cov(y3,y3)\times cov(y4,y4)-cov(y3,y4)\times cov(y4,y3)=0.0126$

  $\Sigma^{-1}=\frac{1}{det(\Sigma)}\begin{bmatrix}cov(y4,y4 & -cov(y3,y4\\ -cov(y4,y3) & cov(y3,y3)\end{bmatrix}=\begin{bmatrix}19.84127 & -14.28571\\ -14.28571 & 14.28571\end{bmatrix}$

  $y3,y4|class=0\sim N(\mu,\Sigma)\sim N\left([0.2\quad 0.25],\begin{bmatrix}0.18 & 0.18\\ 0.18 & 0.25\end{bmatrix}\right)$

- *Class 1*: $\mu=\frac{1}{6}\sum_{i=5}^{10}[y_{3_i}\quad y_{4_i}]=[0.1(6)\quad 0.08(3)]$

  $\Sigma=\begin{bmatrix}cov(y3,y3) & cov(y3,y4)\\ cov(y4,y3) & cov(y4,y4)\end{bmatrix}=\begin{bmatrix}0.109(6) & 0.122(3)\\ 0.122(3) & 0.213(6)\end{bmatrix}$

  $det(\Sigma)=cov(y3,y3)\times cov(y4,y4)-cov(y3,y4)\times cov(y4,y3)=0.0084(6)$

  $\Sigma^{-1}=\frac{1}{det(\Sigma)}\begin{bmatrix}cov(y4,y4 & -cov(y3,y4\\ -cov(y4,y3) & cov(y3,y3)\end{bmatrix}=\begin{bmatrix}25.23622 & -14.44882\\ -14.44882 & 12.95276\end{bmatrix}$

Formatted: Underline
Formatted: No underline
Formatted: Font: Italic
Formatted: No bullets or numbering
Formatted: Font: Italic
Formatted: Underline
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Underline
Formatted: Font: Italic, English (United States)
Formatted: Font: Italic
Formatted: Underline
Formatted: Font: Italic
Formatted: Indent: Left: 0.5", No bullets or
Formatted: Font: Italic
Formatted: Indent: Left: 0.5", No bullets or

Aprendizagem 2021/22

Homework I – Group 013

$$y3, y4|class = 1 \sim N(\mu, \Sigma) \sim N\left(\begin{bmatrix}0.1(6) & 0.08(3)\end{bmatrix}, \begin{bmatrix}0.109(6) & 0.122(3) \\ 0.122(3) & 0.213(6)\end{bmatrix}\right)$$

~~Class 1:~~

2) $P(class = c \mid x) = \frac{P(x \ = [Y1, Y2, Y3, Y4]|\ class \ = c)P(class=c)}{P(x=[Y1,Y2,Y3,Y4])}$

2)

Assuming naïve Bayes : $P(x \ = \ [Y1, Y2, Y3, Y4] \mid class \ = c) \ = \ P(y1 \ = \ Y1| \ class = c) \ P(y2 \ = \ Y2 \mid class = c) \ P(y3 \ = \ Y3, y4 = Y4 \ |class = c)$, where the likelihood of each conditional variable to the class is given by the distributions calculated in question **1)**

Normalization: $\frac{P(Class \ = c|x)}{P(Class \ = c|x)+P(\neg Class \ = c|x)}$

	x1		x2		x3		x4		x5	
Class	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1
P(class = c)	0.4	0.6	0.4	0.6	0.4	0.6	0.4	0.6	0.4	0.6
P(y1=Y1|class=c)	0.569	0.224	1.374	1.364	1.639	1.209	1.374	1.364	1.639	0.950
P(y2=Y2|class=c)	0.5	0.1(6)	0.25	0.(3)	0.5	0.1(6)	0.25	0.5	0.25	0.(3)
P(y3=Y3, y4=Y4 | class = c)	1.207	1.210	0.460	0.955	0.707	0.610	0.512	0.203	1.174	1.206
P(x | class = c)	0.343	0.045	0.158	0.434	0.579	0.123	0.176	0.138	0.481	0.382
P(class = c | x)	1.373	0.271	0.633	2.605	2.317	0.738	0.704	0.829	1.925	2.293
Normalization	0.835	0.165	0.195	0.805	0.758	0.242	0.459	0.541	0.456	0.544

	x6		x7		x8		x9		x10	
Class	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1	c = 0	c = 1
P(class = c)	0.4	0.6	0.4	0.6	0.4	0.6	0.4	0.6	0.4	0.6
P(y1=Y1|class=c)	0.569	1.209	0.116	0.662	1.639	1.209	1.374	0.662	0.280	0.950
P(y2=Y2|class=c)	0.25	0.5	0.25	0.5	0.25	0.(3)	0.5	0.1(6)	0.25	0.5
P(y3=Y3, y4=Y4 | class = c)	0.334	0.672	0.707	0.610	1.085	0.838	0.217	0.387	1.080	1.123

Formatted: Indent: Left: 0", First line: 0"

Formatted: Indent: First line: 0"

Formatted: Indent: Left: -0.25", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

Formatted: Font: Bold

Formatted Table

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted Table

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

TÉCNICO LISBOA

Aprendizagem 2021/22
Homework I – Group 013

P(x | class = c)	0.047	0.406	0.021	0.202	0.445	0.338	0.149	0.043	0.076	0.534
P(class = c | x)	0.190	2.437	0.082	1.212	1.778	2.027	0.598	0.256	0.303	3.202
Normalization	0.072	0.928	0.063	0.937	0.467	0.533	0.700	0.300	0.087	0.913

$$true \ = < 0,0,0,0,1,1,1,1,1,1 > \ and\ predicted = < 0,1,0,1,1,1,1,1,0,1 >$$

	0 (predicted)	1 (predicted)	Sum
0 (true)	2	2	4
1 (true)	1	5	6
Sum	3	7	10

~~x1~~ ~~$P(class = 0|x1) =$~~

~~P(class = 1|x1) =~~

~~x2~~

~~x3~~

~~x4~~

~~x5~~

~~x6~~

~~x7~~

~~x8~~

~~x9~~

~~x10~~

3) ~~Answer 4~~Class 0: $Precision \ = \frac{True\ Positives}{True\ Positives + False\ Positives} = 0.(6)$ $Recall \ = \frac{True\ Positives}{Positives} = 0.5$

$$\frac{1}{F1} = \frac{1}{2}\left(\frac{1}{Recall} + \frac{1}{Precision}\right) \Leftrightarrow F1 \ = \ 0.(36)$$

Class 1: $Precision \ = \frac{True\ Positives}{True\ Positives + False\ Positives} = 0.714285$ $Recall \ = \frac{True\ Positives}{Positives} = 0.8(3)$

$$\frac{1}{F1} = \frac{1}{2}\left(\frac{1}{Recall} + \frac{1}{Precision}\right) \Leftrightarrow F1 \ = \ 0.526316$$

4) We used the posteriors as thresholds, because between them, the obtained results from the classifier are the same. So, we can conclude that the best threshold to use is t_{best} where $t_{best} \in$ (0.70014203,0.758462001)~~daskdlçaskdç~~

True	P(class=0|x)	0.06344	0.07224	0.08653	0.19538	0.45643	0.45923	0.46736	0.70014	0.75846	0.8352
0	0.835224838	0	0	0	0	0	0	0	0	0	1
0	0.195377549	0	0	0	1	1	1	1	1	1	1
0	0.758462001	0	0	0	0	0	0	0	0	1	1
0	0.459234927	0	0	0	0	0	1	1	1	1	1
1	0.456434993	0	0	0	0	1	1	1	1	1	1

Formatted: Centered
Formatted: Centered
Formatted: Centered
Formatted: Font: (Default) Times New Roman
Formatted Table
Formatted: Font: 4 pt
Formatted: Normal, Indent: First line: 0"
Formatted: Indent: Left: -0.25", Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Bar : (No border)
Formatted: No bullets or numbering
Formatted: No bullets or numbering
Formatted: No bullets or numbering, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Bar : (No border)
Formatted: Indent: Left: -0.25"
Formatted: English (United States)
Formatted: Left
Formatted: Centered
Formatted: Centered
Formatted: Centered
Formatted: Centered
Formatted: Centered

Aprendizagem 2021/22

Homework I – Group 013

1	0.072240527	0	1	1	1	1	1	1	1	1	1
1	0.063438589	1	1	1	1	1	1	1	1	1	1
1	0.467357063	0	0	0	0	0	0	1	1	1	1
1	0.700142029	0	0	0	0	0	0	0	1	1	1
1	0.086525992	0	0	1	1	1	1	1	1	1	1
accuracy		0.5	0.6	0.7	0.6	0.7	0.6	0.7	0.8	0.7	0.6

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

Formatted: Centered

II. Programming and critical analysis

5) Answer 5

6) Answer 6

7) Answer 7

8) Answer 8

III. APPENDIX

```
Paste your programming code here using Consolas 9pt or 10pt.
Use highlighting or colored text to facilitate the analysis by your faculty hosts.
```

END