Exercise 3: Feature Processing

Pablo Castro Urandain PR

1

- What does a ROC curve represent? → It's the representation of the ratio of True positives against the ratio of False positives. It represents the sensibility in contrast to the ~~the~~ specification (or in contrast to) ~~~~ of a binary classifier ~~~~ as soon as the threshold of discrimination is changing his value

- How can the ROC curve be produced for a binary classification case? → First, we would have to distinguish our binary data into 4 categories: True Positives, False Positives, True Negatives; False negatives

Real value

		p	n	total
Prediction Value	p'	TP	FP	P'
	n'	FN	TN	N'
	total	P	N	

→ Contingency table

Supposing we have all our data divided, now we should have a threshold for producing our curve. The curve is gona look like the line that passes through all of the threshold points. So if the value of our threshold is 0,7, for example, we can say that we have Positives if the value surpasses (the value of our probability p). If not, ~~the~~ we have a Negative. The curve is made out of all the different ~~~~ outputs of our data when stablised a threshold, so we are going to be creating different contingency tables with our different thresholds. For example: 0,1, 0,2, 0,3, 0,4, ...

- How is it possible to use the ROC curve to determine the optimal threshold for a binary classifier?

→ By training our system with lots of data, the threshold won't be the perfect one, but as much as we keep it training, the ~~~~ & output threshold is going to be the best for our system

(Changing the value of our threshold and look for the best one)

②

$$FDR_1 = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

Feature 1 $\rightarrow$ $\mu_1 = 3 \quad \sigma_1 = 2$, $\mu_2 = 7 \quad \sigma_2 = 1$ $\Big\}$ $FDR_1 = \frac{(3-7)^2}{2^2+1^2} = \frac{16}{5} = 3,2$

Feature 2 $\rightarrow$ $\mu_1 = 5$, $\mu_2 = 6$, $\sigma_1 = \sigma_2 = 0,2$ $\Big\}$ $FDR_1 = \frac{(5-6)^2}{0,2^2+0,2^2} = \frac{1}{0,08} = 12,5$ // $\rightarrow$ $\boxed{\text{Better}}$

④

$X_1 = \begin{pmatrix} 8 \\ 9 \end{pmatrix} \quad X_2 = \begin{pmatrix} 1 \\ 9 \end{pmatrix} \quad X_3 = \begin{pmatrix} 6 \\ 1 \end{pmatrix} \quad X_4 = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad X_5 = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$

a) Matrix with both features: (Min - Max Normalisation)

$$\begin{pmatrix} 8 & 1 & 6 & 3 & 10 \\ 9 & 9 & 1 & 5 & 10 \end{pmatrix}$$

$\forall \quad \hat{X}_{ik} = \frac{X_{ik} - X_k^{min}}{X_k^{max} - X_k^{min}} \Rightarrow$

$\hat{X}_{11} = \frac{8-1}{10-1} = \frac{7}{9} = 0,77$ | $\hat{X}_{21} = \frac{9-1}{10-1} = \frac{8}{9} = 0,88$

$\hat{X}_{12} = \frac{1-1}{10-1} = 0$ | $\hat{X}_{22} = \frac{9-1}{10-1} = \frac{8}{9} = 0,88$

$\hat{X}_{13} = \frac{6-1}{10-1} = \frac{5}{9} = 0,55$ | $\hat{X}_{23} = \frac{1-1}{10-1} = 0$

$\hat{X}_{14} = \frac{3-1}{10-1} = \frac{2}{9} = 0,22$ | $\hat{X}_{24} = \frac{5-1}{10-1} = \frac{4}{9} = 0,44$

$\hat{X}_{15} = \frac{10-1}{10-1} = \frac{9}{9} = 1$ | $\hat{X}_{25} = \frac{10-1}{10-1} = \frac{9}{9} = 1$

• Result Matrix:

$$\begin{pmatrix} 0,77 & 0 & 0,55 & 0,22 & 1 \\ 0,88 & 0,88 & 0 & 0,44 & 1 \end{pmatrix}$$

b) Mean - Variance Normalisation:

$$\hat{X}_{ik} = \frac{X_{ik} - \bar{X}_k}{\sigma_k} \Rightarrow \bar{X}_1 = \frac{8+1+6+3+10}{5} = \frac{28}{5} = 5{,}6$$

$$\bar{X}_k = \frac{1}{N}\sum_{i=1}^{N} X_{ik} \qquad \bar{X}_2 = \frac{9+9+1+5+10}{5} = \frac{34}{5} = 6{,}8$$

$$\sigma_k^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_{ik} - \bar{X}_k\right)^2 \Rightarrow \sigma_1 = \sqrt{\frac{(8-5{,}6)^2 + (1-5{,}6)^2 + (6-5{,}6)^2 + (3-5{,}6)^2 + (10-5{,}6)^2}{5-1}} = 3{,}64$$

$$\sigma_2 = \sqrt{\frac{(9-6{,}8)^2 + (9-6{,}8)^2 + (1-6{,}8)^2 + (5-6{,}8)^2 + (10-6{,}8)^2}{5-1}} = 3{,}76$$

$$\hat{X}_{11} = \frac{8-5{,}6}{3{,}64} = 0{,}65 \qquad \hat{X}_{21} = \frac{9-6{,}8}{3{,}76} = 0{,}58$$

$$\hat{X}_{12} = \frac{1-5{,}6}{3{,}64} = -1{,}26 \qquad \hat{X}_{22} = \frac{9-6{,}8}{3{,}76} = 0{,}58$$

$$\hat{X}_{13} = \frac{6-5{,}6}{3{,}64} = 0{,}10 \qquad \hat{X}_{23} = \frac{1-6{,}8}{3{,}76} = -1{,}53$$

$$\hat{X}_{14} = \frac{3-5{,}6}{3{,}64} = -0{,}7 \qquad \hat{X}_{24} = \frac{5-6{,}8}{3{,}76} = -0{,}49$$

$$\hat{X}_{15} = \frac{10-5{,}6}{3{,}64} = 1{,}21 \qquad \hat{X}_{25} = \frac{10-6{,}8}{3{,}76} = 0{,}85$$

Result Matrix:

$$\begin{pmatrix} 0{,}65 & -1{,}26 & 0{,}1 & -0{,}7 & 1{,}21 \\ 0{,}58 & 0{,}58 & -1{,}53 & -0{,}49 & 0{,}85 \end{pmatrix}$$

c) Softmax - Scaling:

$$\hat{X}_{ik} = \frac{1}{1+e^{-Y_{ik}}} \Rightarrow \hat{X}_{11} = \frac{1}{1+e^{-0{,}65}} = 0{,}66 \qquad \hat{X}_{21} = \frac{1}{1+e^{-0{,}58}} = 0{,}64$$

$$\hat{X}_{12} = \frac{1}{1+e^{1{,}26}} = 0{,}22 \qquad \hat{X}_{22} = \frac{1}{1+e^{-0{,}58}} = 0{,}64$$

$$Y_{ik} = \frac{X_{ik} - \bar{X}_k}{\sigma_k} \qquad \hat{X}_{13} = \frac{1}{1+e^{0{,}1}} = 0{,}54 \qquad \hat{X}_{23} = \frac{1}{1+e^{1{,}53}} = 0{,}19$$

$$\hat{X}_{14} = \frac{1}{1+e^{0{,}7}} = 0{,}32 \qquad \hat{X}_{24} = \frac{1}{1+e^{0{,}49}} = 0{,}38$$

(on the previous ~~a~~ b), it was the $\hat{X}_{ik}$)

$$\hat{X}_{15} = \frac{1}{1+e^{-1{,}21}} = 0{,}77 \qquad \hat{X}_{25} = \frac{1}{1+e^{-0{,}85}} = 0{,}7$$

$$\begin{pmatrix} 0{,}66 & 0{,}22 & 0{,}54 & 0{,}32 & 0{,}77 \\ 0{,}64 & 0{,}64 & 0{,}19 & 0{,}38 & 0{,}7 \end{pmatrix} \Rightarrow \text{Result Matrix}$$