Section 4L. Confusion Matrix

Statistics for Data Science

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Confusion Matrix

The ***confusion matrix*** is defined as

	$y_i = 0$	$y_i = 1$	
$C(\mathbf{x}_i) = 0$	$\lvert\mathsf{TN}\rvert$	$\lvert\mathsf{FN}\rvert$	$\lvert\mathsf{TN}\rvert + \lvert\mathsf{FN}\rvert$
$C(\mathbf{x}_i) = 1$	$\lvert\mathsf{FP}\rvert$	$\lvert\mathsf{TP}\rvert$	$\lvert\mathsf{FP}\rvert + \lvert\mathsf{TP}\rvert$
	$\lvert\mathsf{TN}\rvert + \lvert\mathsf{FP}\rvert$	$\lvert\mathsf{FN}\rvert + \lvert\mathsf{TP}\rvert$	N

Confusion Matrix: LDA Practice

We train an LDA classifier $C_{\text{LDA}}(x)$ using the Default dataset and show results below:

- The confusion matrix takes the form

	$y_i = 0$	$y_i = 1$	
$C(\mathbf{x}_i) = 0$	9644	252	9896
$C(\mathbf{x}_i) = 1$	23	81	104
	9667	333	10000

- The classification error rate (in the training dataset) is

$$\text{Err}_{\text{LDA}} = \frac{23 + 252}{10000} = 2.75\%$$

- Is this a good classification rate? For example, the *null classifier* $C_{\text{Null}}(x_i) = 0$ has $|FP| = 0$ and $|FN| = 333$; hence,

$$\text{Err}_{\text{Null}} = \frac{333}{10000} = 3.33\%$$

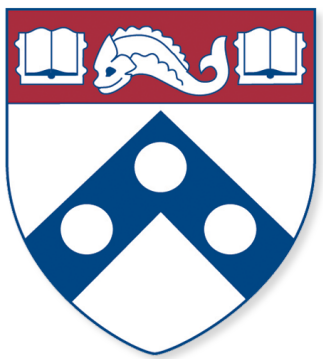
Classification Error Rates

Observations on the LDA classifier

- For those samples for which $y_i = 0$, we make an error rate of $23/9667 = 0.2\%$
- However, for those samples for which $y_i = 1$, our error rate is $252/333 = 75.7\%$

Hence, the classification error rate alone is not a good measure of performance...

- We can define more refined versions of the classification rate:
  - *True Positive Rate*: $\text{TPR} = |\text{TP}| / (|\text{TP}| + |\text{FN}|)$
  - *True Negative* Rate*:* $TNR = |TN| / (|TN| + |FP|)$
  - *False Positive* Rate*:* $FPR = |FP| / (|FP| + |TN|) = 1 - TNR$
  - *False Negative* Rate*:* $FNR = |FN| / (|FN| + |TP|) = 1 - TPR$



Penn
Engineering