

Classification Metrics

Damien Benveniste

June 17, 2017

1 The confusion matrix

Here is the confusion matrix

		predicted condition	
		prediction positive	prediction negative
true condition	condition positive	True Positive (TP)	False Negative (FN) (type II error)
	condition negative	False Positive (FP) (Type I error)	True Negative (TN)

There are few important quantity to know:

- The accuracy:

$$a = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \quad (1)$$

- The misclassification error:

$$\epsilon = 1 - a \quad (2)$$

- The true positive rate:

$$TPR = \frac{\sum TP}{\sum TP + \sum FN} \quad (3)$$

- The false positive rate:

$$FPR = \frac{\sum FP}{\sum FP + \sum TN} \quad (4)$$

2 Receiver operating characteristic curve

The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

