

Confusion matrix

Recall = $\frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{91}{49 + 91} = \frac{91}{140}$

Correctly predicted

Precision = $\frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{91}{41 + 91} = \frac{91}{132}$

Actual \ Preds	Preds	
	0	1
0	212	41
1	49	91

FP (False Positive) is 41, TP (True Positive) is 91.

$F_1 = \frac{2PR}{P+2}$ (Harmonic mean)

Smallest mean

Am: Values are high it will be high

n=165

	Predicted	
	0	1
0	50	10
1	5	100

	Actual	
	TN	FP
0		
1		

FN (False Negative) is 5, TP (True Positive) is 100.

- There are two possible predicted classes yes and no
- If we are predicting a disease, for ex: "yes" would mean they have the disease and "no" would mean they don't have the disease.
- Out of 165 cases, the classifier predicted "yes" 110 times and "no" 55 times
- In reality 105 in the sample have the disease and 60 patients do not.
- True positive: There are cases in which we predicted yes they do have the disease
- True negative: There are cases in which we predicted no they don't have the disease.
- False positive: We predicted yes, but they don't actually have disease (Type I Error)
- False negative: We predicted no, but they have the disease. (Type II Error)

Accuracy = $\frac{\text{TP} + \text{TN}}{\text{Total}} = \frac{100 + 50}{165} = 0.91$

misclassification rate (How often it is wrong) = $\frac{\text{FN} + \text{FP}}{\text{Total}} = \frac{10 + 5}{165} = 0.09$

(Recall)

True Positive Rate \div when it actually yes, how often it predicted yes

$$TP/Actual\ yes = \frac{TP}{TP+FN} = \frac{100}{105} = 0.95$$

False positive rate \div when it actually no, how often it predicted yes

$$FP/actual\ no \div \frac{FP}{TN+FP} = 10/60 = 0.17$$

Specificity \div when it is actually no, how often it predicted no

$$\frac{TN}{TN+FP} = 50/60 = 0.83$$

Precision \div when it ~~is~~ predicts yes, how often it is correct

$$\frac{TP}{FP+TP} = \frac{100}{110} = 0.91$$

Prevalance \div How often yes condition occur in the sample.

$$actual\ yes / total = \frac{105}{165} = 0.64$$

	Predict	
	LT 1	NLT 0
ALT 1	A	B
ANLT 0	C	D
	Actual	

$$Recall = \frac{A}{A+B}$$

$$Precision = \frac{A}{A+C}$$

$$F_1 = \frac{2PR}{P+R}$$

	Correct	Incorrect
Actual cancer	A	B
Actual no cancer	C	D

B is very important when we are predicting for free

\rightarrow B & C is very important when we are predicting for cost.

Score.