

Performance Metrics (Binary Classification)

- | | |
|----------------------------|--|
| 1. Confusion Matrix | 1 - Positive \rightarrow True |
| 2. Accuracy | 0 - Negative \rightarrow False |
| 3. Precision | |
| 4. Recall | |
| 5. F-Beta Score (F1 Score) | $1, 1 \rightarrow$ True Positive
$0, 1 \rightarrow$ False Positive
$0, 0 \rightarrow$ True Negative
$1, 0 \rightarrow$ False Negative |

dataset (Features $\rightarrow x_1, x_2, \dots, y, \hat{y}$)

Actual			
x_1	x_2	y	\hat{y} \leftarrow Predicted Value
-	-	1	0 \leftarrow FN
-	-	1	1 \leftarrow TP
		0	1 \leftarrow FP
		0	0 \leftarrow TN
		0	1 \leftarrow FP
		1	1 \leftarrow TP
		1	1 \leftarrow TP
		1	0 \leftarrow FN

		Actual	
		1	0
Predicted	1	TP 3	FP 2
	0	FN 2	TN 1

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\Rightarrow \frac{3 + 1}{3 + 1 + 2 + 2} \Rightarrow \frac{4}{8} = 0.5 = 50\%$$

- Spam
- ① $\left(\begin{array}{cc} \text{Actual} & \text{Predicted} \\ \text{Spam} & \text{Spam} \end{array} \right) \leftarrow$ TP
- ② $\left(\begin{array}{cc} \text{Spam} & \text{NotSpam} \\ \text{NotSpam} & \text{Spam} \end{array} \right) \leftarrow$ FN
- ③ $\left(\begin{array}{cc} \text{NotSpam} & \text{Spam} \end{array} \right) \leftarrow$ FP
- ④ $\left(\begin{array}{cc} \text{NotSpam} & \text{NotSpam} \end{array} \right) \leftarrow$ TN

Precision

10,000

80% \rightarrow T \rightarrow 8000

2000 \rightarrow F

$$= \frac{TP}{TP + FP} \downarrow$$

$$\text{Recall} = \frac{TP}{TP + FN} \downarrow$$

F Beta Score

$$= (1 + \beta^2) \frac{\text{Precision} * \text{Recall}}{\beta^2 * (\text{Precision} + \text{Recall})}$$

- ① False Positive and FN \leftarrow Important $\leftarrow \beta = 1 \leftarrow (1+1) \frac{P * R}{(1)(P+R)}$
- ② F. Positive $>$ FN $\leftarrow \beta = 0.5 \leftarrow 1 + (0.25) \frac{P * R}{(0.25)(P+R)}$
- ③ FN \gg FP $\leftarrow \beta = 2 \leftarrow (1+4) \frac{P * R}{(4)(P+R)}$