→ Evaluation of a Classification model

Target y	Prediction $\hat{y}$
A	B
B	B
→ (A)	(A)
(A)	(B)
A	B
(B)	(A)
(B)	(B)

(4) → 4 x 4

(A - B)

Subtraction

→ N x N → # of categories

(1) Matrix → Confusion Matrix

Classification Matrix

Organizing the comparison

PRED \ Actual	A	B
A	TP	FP
B	FN	TN

How good the prediction → Evaluation for Regression Models

Performance

Quality

↳ How well the model is performing.

- → R^2 → strength overall
- → MSE → Mag. of Error
- → RMSE →
- → Adj R^2 → strength w.r.t error
- → VIF → multicollinearity
- → MAE → magnitude of error

Target y	Prediction $\hat{y}$
1	0.7
2	0.8
3	0.9
4	3.14

Continuous

Comparing (y, $\hat{y}$)

⇓

Difference (−)

$Err = y - \hat{y}$

$1 - 0.7$

$= 0.3$

Pred \ Actual	+ve	-ve
+ve	TP	FP (→ α)
-ve	FN (→ β)	TN

T	P
If the Pred is correct or not	what is the Prediction class

A , B
↓ ↓
+ve -ve

Success / Failure

Buy / Not Buy

Cancer / Not cancer

Spam / Not Spam

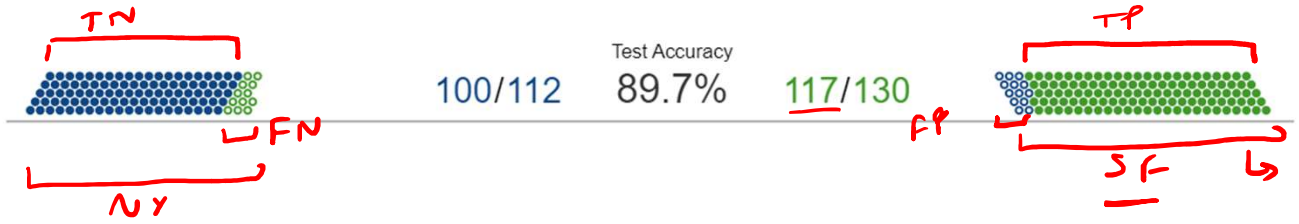
TP ⇒ The # of Rows that are predicted as +ve & are actually +ve ✓

FP ⇒ (→ α-errors) " " " " " +ve " " -ve } → Errors

FN ⇒ (→ β-errors) " " " " " -ve " " +ve

TN ⇒ " " " " " -ve " " -ve ✓

Nothing! → This step is only to organize the results
Using the matrix we will evaluate the perf.



Blue → NY → -ve

Green → SF → +ve

TP : 117 ✓

TN : 100 ✓

FN : 12 ✗

FP : 13 ✗

A	+	-
+	117	13
-	12	100

① Accuracy → $\frac{\text{Total correct Pred.}}{\text{Total Pred.}} = \frac{TP + TN}{TP + FP + FN + TN} = \frac{217}{242} = 89.7\%$

② Error → $\frac{\text{Total incorrect Pred.}}{\text{Total Pred.}} = \frac{FN + FP}{TP + FP + FN + TN} = 1 - \text{Accuracy} = \frac{25}{242} = 10.3\%$

↳ Overall Perf.

↳ We also need Class-wise Perf.

P \ A	+ve	-ve
+ve	800	200
-ve	0	0

800 / 1000

Acc = 80%

Err → 20%

→ Rows → 1000 → 800 +ve, 200 -ve

give the output as +ve → 1000 +ve

⇒ Recall (-ve) ⇒ 0

Precision (-ve) ⇐ 0

→ Class wise Performance

(1) Precision → Out of total Class Pred., How many are correct

Pred

+ve -ve -ve

↳ Out of total +ve Pred, How many are correct!

$$\text{Precision (+ve)} \Rightarrow \frac{TP}{TP+FP} = \frac{117}{130} = 90\%$$

$$\text{Precision (-ve)} \Rightarrow \frac{TN}{FN+TN} = \frac{100}{112} = 89.3\%$$

Row (Actual)

(2) Recall → out of total Actual Class, How many are Predicted Correctly?

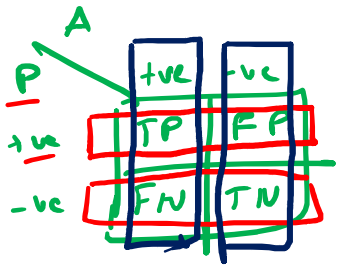
(+ve) (-ve)

↳ out of total Actual +ve, How many are Predicted Correctly?

$$\text{Recall (+ve)} = \frac{TP}{TP+FN} = \frac{117}{129} \Rightarrow 90.7\%$$

$$\text{Recall (-ve)} = \frac{TN}{FP+TN} = \frac{100}{113} = 88.5\%$$

11 class
11
11 Pairs



117	13
12	100

→ Humans ⇒ lazy

F_1 = 60%

F_1 = 84%

Acc = 80%
F_1 = 85%

(5) → F1 - Score → Harmonic Mean of Precision & Recall

Arithmetic Mean (P & R) = $\frac{P+R}{2}$

Harmonic mean (P & R) = $\frac{2}{\frac{1}{P}+\frac{1}{R}} = \frac{2PR}{P+R}$

Cases	I	II	III	IV
	P = 10% | R = 90%	P = 90% | R = 10%	P = 10% | R = 15%	P = 90% | R = 80%
AM ⇒	50	50	12.5	85
HM ⇒	18	18	12	84.7

Max

Min

↳ when both P & R are high

Either Both or any one is low (P & R) (P or R) then, F1-Score is low

→ Quality → ROC Curve (Receiver Operating Characteristic Curve) → AUC (ROC)

⇒ TPR vs FPR

⇒ TPR (TP Rate) = $\frac{TP}{TP + FN \text{ (Actual +ve)}}$

⇒ FPR (FP Rate) ↳ = $\frac{FP}{FP + TN \text{ (Actual -ve)}}$

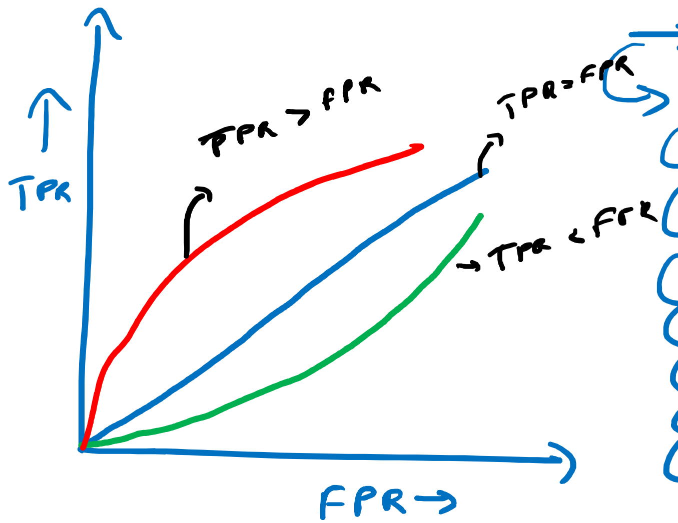
TPR = FPR ⇒ AUC = 50% bad

TPR < FPR ⇒ AUC < 50% bad

TPR > FPR ⇒ AUC > 50% → good model

AUC ≥ 75% → very good model

[−∞, +∞] → [0, 1] → Prob. → P3 → 1 | 0 → ≥ 0.7



Row #	Actual Class	Prob.	TP	FP	TPR	FPR
1	P.	0.9 P	1	0	1/4 = 0.25	0/4 = 0
2	P.	0.8 P	2	0	0.5	0
3	N	0.7 P	2	1	0.5	0.25
4	P.	0.4 N	2	1	0.5	0.25
5	N	0.9 P	2	2	0.5	0.5
6	N	0.8 P	2	3	0.5	0.75
7	N	0.3 N	2	3	0.5	0.75
8	P	0.7 P	3	3	0.75	0.75
⋮	⋮	⋮	⋮	⋮	⋮	⋮