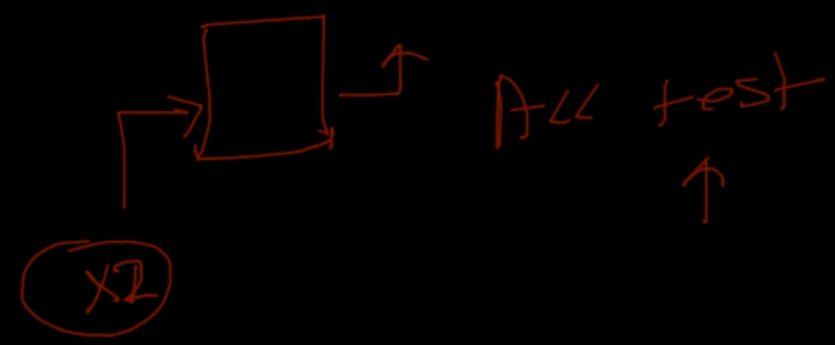
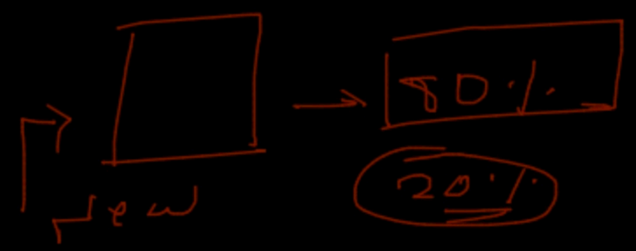
$$\frac{\#\text{ correctly classified pts}}{\text{total } \#\text{ of pts in } D_{test}}$$

eg:

	Actual
D_{test}	
X1	Y ✗
X2	O O
X3	1 O
X4	1 1
X5	1 1

$\frac{4}{5} \rightarrow 80\%$





Acc
10000
D
Dtrain (Xtrain, Ytrain)
Dtest (Xtest, Y-test)
2000
classification
Yes
NO
Y → 0 → 90%
1 → 10%
f( )
good
50%
Best
X
Y
(Xtrain, Ytrain)
0 → 90%
1 → 10%
2000
1500 (Yes)
200 (NO)
(Best)
70%
overall
1800
1500
(300)
200
0%
Acc
10%

→ imbalance → X Acc

(1)

X	Y	M1	M2	$\hat{Y}_1$	$\hat{Y}_2$
x1	1	0.9	0.6	1	1 ✓
x2	1	0.8	0.65	1	1
x3	0	0.1	0.45	0	0
x4	0	0.15	0.48	0	0

100 → M1 (Best) → M1 → yes/No

Rec → M2() → M2 → yes/No

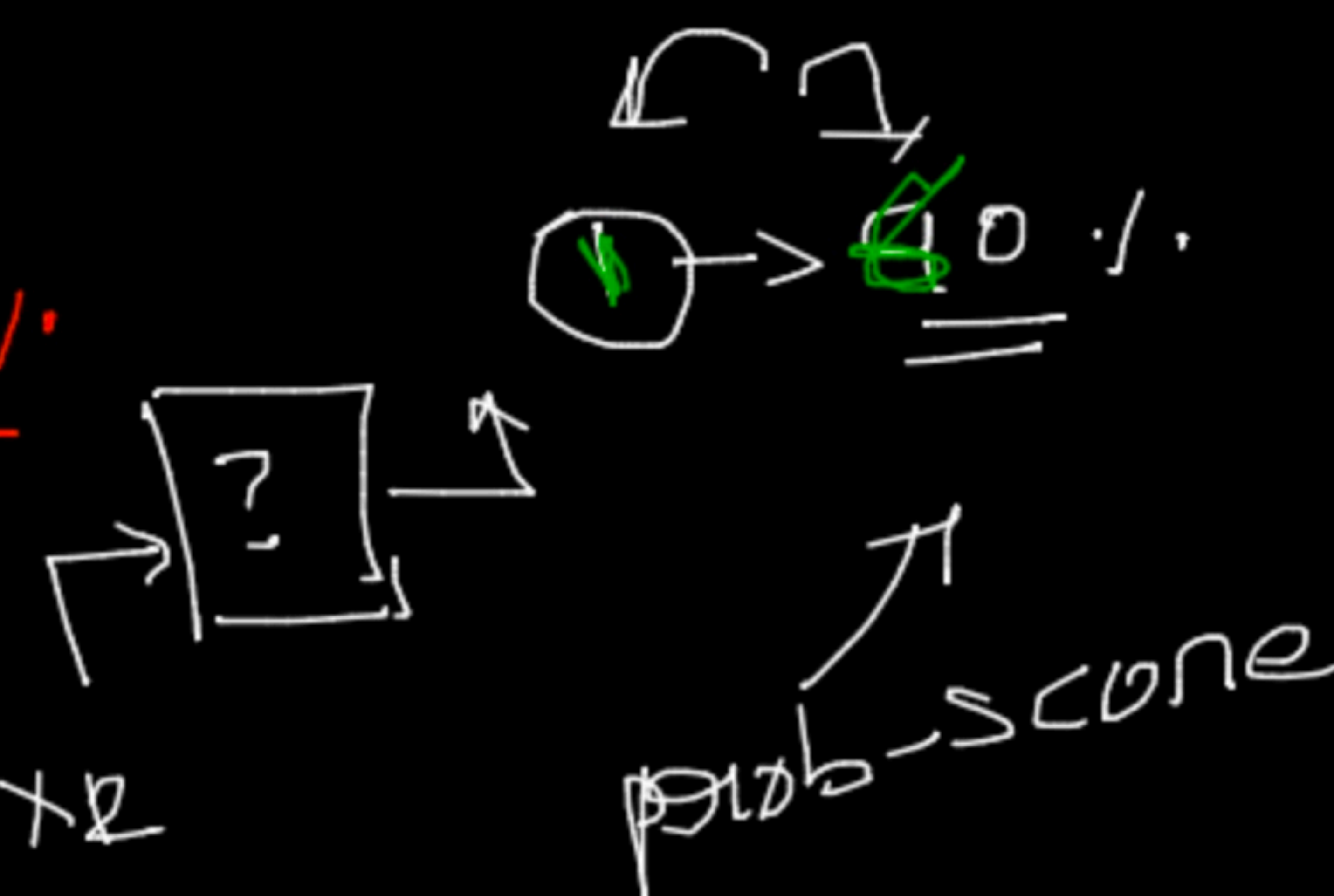
if score > 0.5 → class

else → 0

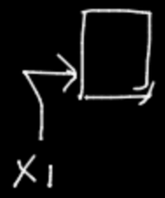
Acc(M2) = 4/4 → 100%

Acc(M1) = 4/4 → 100%

KPI



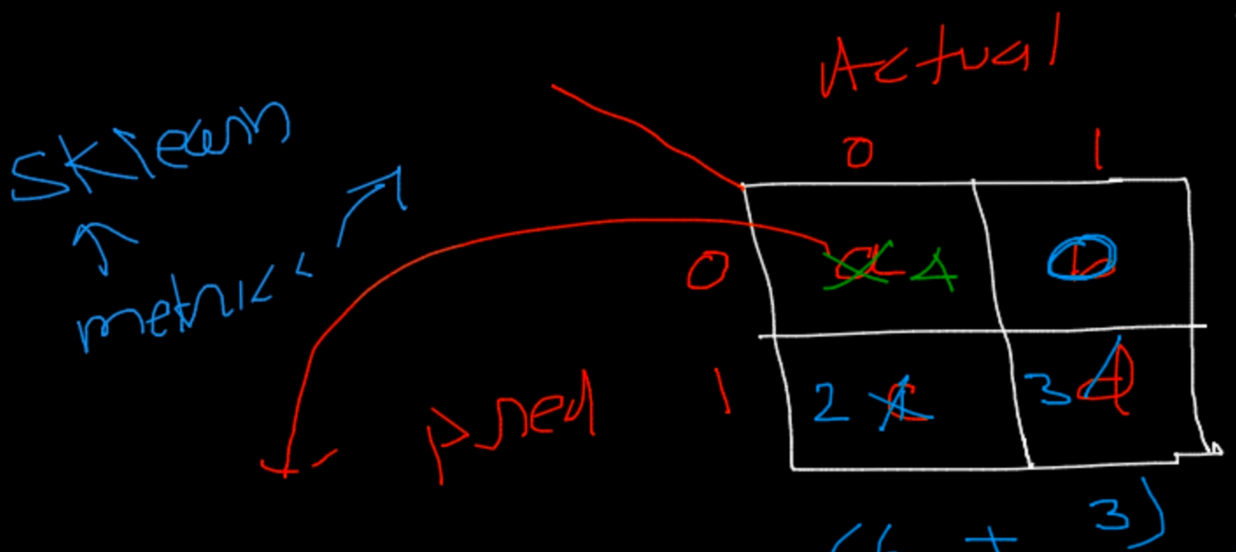
(II) confusion matrix



testing (Yes / No)

eg.

	Y_actual	Y
x1	0	0 ✓
x2	0	0 ✓
x3	0	1 ✗
x4	1	1 ✓
x5	1	1 ✓
x6	0	0 ✓
x7	1	1 ✓
x8	0	1 ✗
x9	0	0 ✓



$q \leftarrow (6 + 3)$

a: # points such that

$y_i = 0$

$\hat{y}_i = 0$

(good model)

$b \rightarrow \neq b$


Actual
0
1
prediction
0
1
TN
FN
FP
TP
(N + P)
total #data pts
FP
P
2
3
TPR
T P
what is predicted label
are you correct or not

(i) TPR (True +ve Rate) = $\frac{\#TP}{P}$

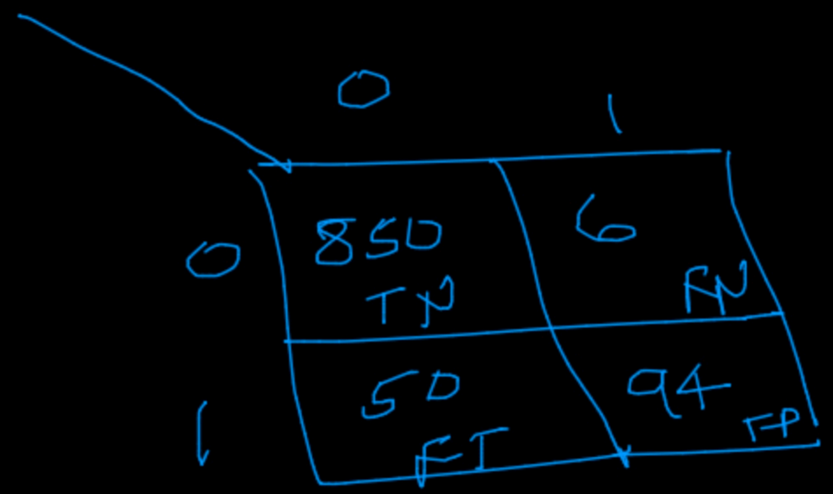
(ii) TNR = $\frac{\#TN}{N}$

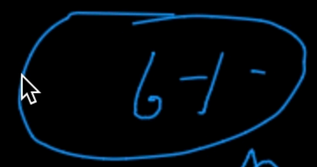
(iii) FPR = $\frac{FP}{P}$ any good

(iv) FNR = $\frac{FN}{N}$

TPR ↑ TNR ↑

and FPR ↓ FNR ↓





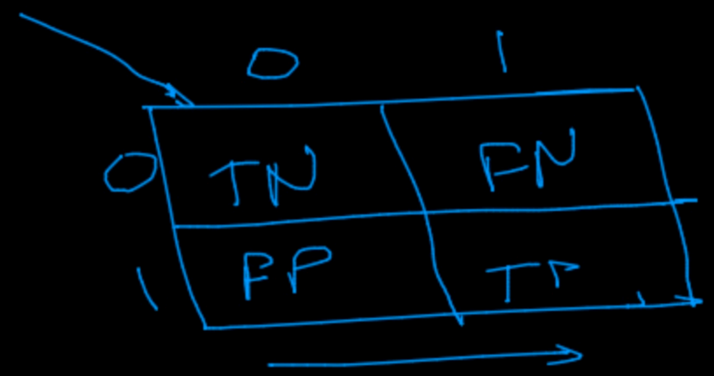
$$FNR = \frac{6}{100}$$

$$TPR = \frac{94}{100} \rightarrow 94\%$$

$$TNR = \frac{850}{900} \rightarrow ?$$

$$FPR = \frac{50}{900}$$

Precision, Recall and F1-score



Recall = TPR

$$= \frac{TP}{P}$$

Pr

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

of all pts the model declared | predicted to be +ve, what fraction of them are actually +ve

F1-score ↑

$$= 2 * \left(\frac{Pr \times Re}{Pr + Recall} \right)$$

Good

Pr ↑ Recall ↑ F-score ↑

↳ 1

(0 - 1)