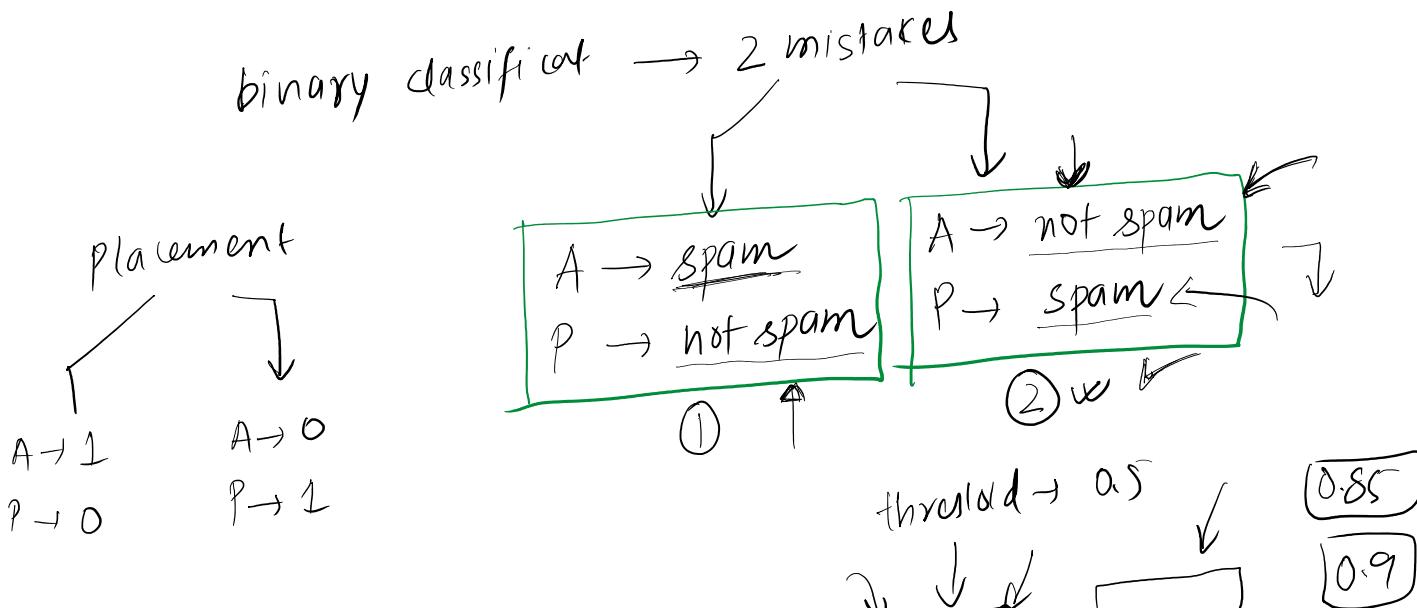
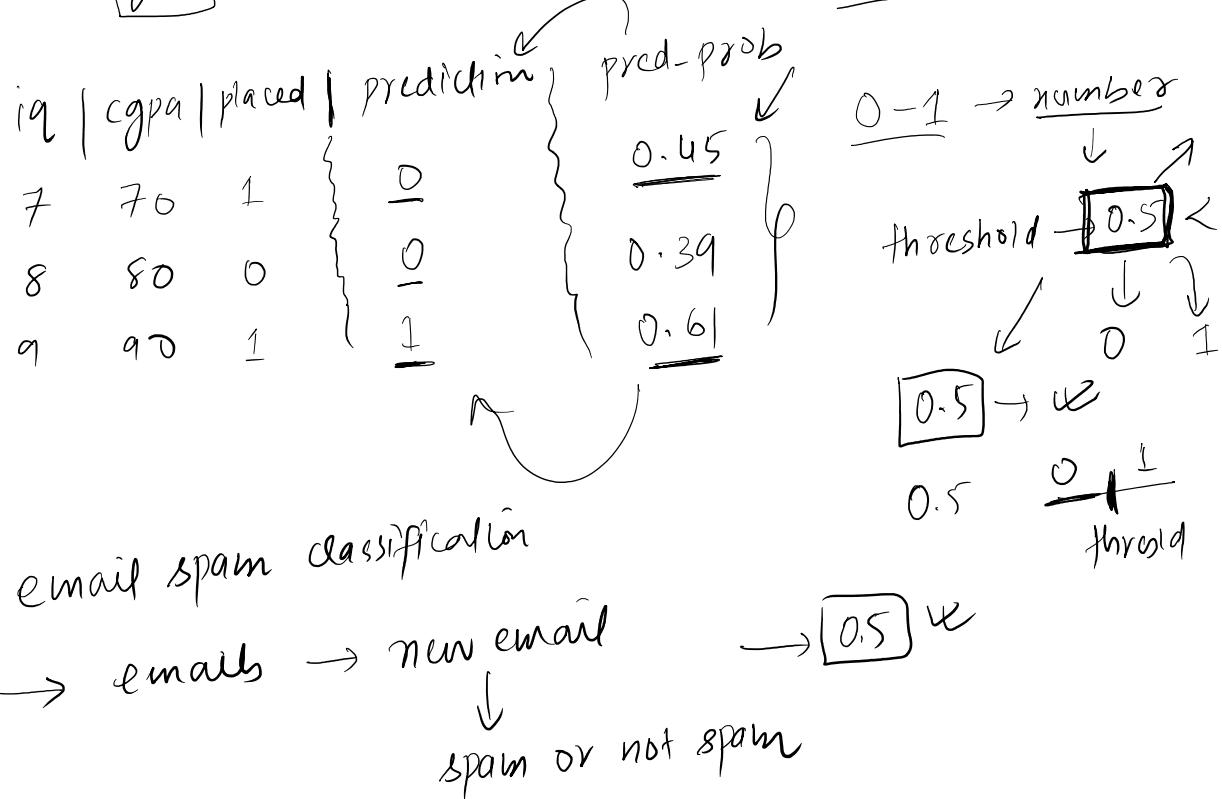
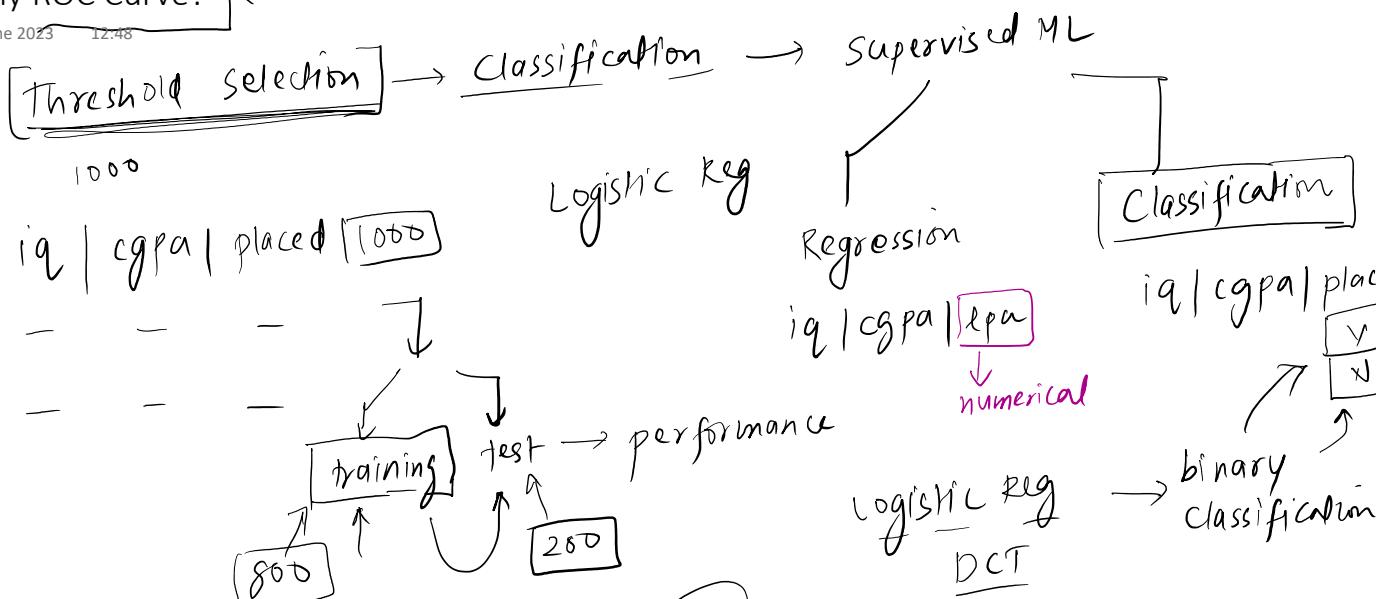
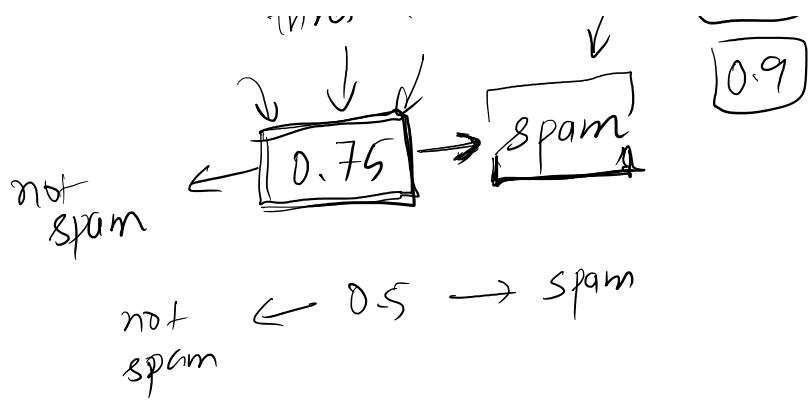
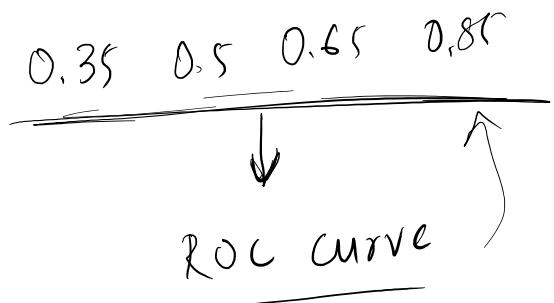


Why ROC Curve?

06 June 2023 12:48



$r \rightarrow 0$



Confusion Matrix

06 June 2023 12:48

		grid		Correct		Predicted	
		Predicted		1000	800	1	0
Actual	1	True Positive	False Negative	200	placed	pred	
	0	False Positive	True Negative	70	1	1	1
		70	8	80	0	15	80
		90	9	1	0	100	5

Report card → binary classification

True Positive Rate (TPR) \rightarrow Benefit $\leftarrow \underline{100\%}$

06 June 2023 12:48

	1	0
1	TP	FN
0	FP	TN

Prediction

email spam classifier

1 \rightarrow spam
0 \rightarrow not spam

$\leftarrow \underline{100\%}$

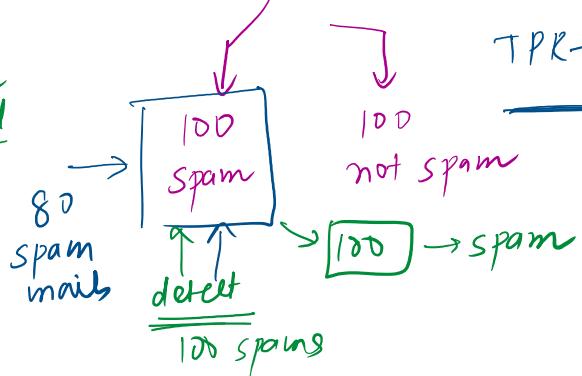
$$\underline{\text{TPR}} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Spam		not spam	
Spam	80	20	not spam
not spam	20	80	

$$\text{TPR} = \frac{80}{80+20} = \frac{80}{100}$$

TPR \rightarrow 80%

test \rightarrow 200 emails



churn rate predict

\rightarrow Netflix

\rightarrow 100 customers

leave

\downarrow

detect

80	20
----	----

$$\text{TPR} = 80\%$$

	1	0
1	TP	FN
0	FP	TN

1 \rightarrow leave
0 \rightarrow not leave

False Positive Rate (FPR) \rightarrow Cost model

06 June 2023 12:49

		1	0
Actual	1	TP	<u>FN</u>
	0	<u>FP = 0</u>	TN = 0
Prediction			

email spam \rightarrow

churn predict \rightarrow

$$\underline{FPR} = \frac{FP}{FP + TN}$$

email spam classifier

100 \rightarrow not spam
 \hookrightarrow spam not

not spam \rightarrow spam

Who were not leaving the platform \rightarrow they will leave

$$\underline{FPR} = \frac{FP}{FP + TN}$$

1 \rightarrow 100%
TN \rightarrow 0

0 \leftarrow
FP = 0

platforms \rightarrow user

100 not spam \rightarrow 20 \leftarrow
 \hookrightarrow spam \leftarrow

20
100

100 \rightarrow 10 \rightarrow
 \hookrightarrow discount

The best case

$$\left\{ \begin{array}{l} TPR = 100\% \text{ or } 1 \\ FPR = 0\% \text{ or } 0 \end{array} \right.$$

best case

FN = 0
FP = 0

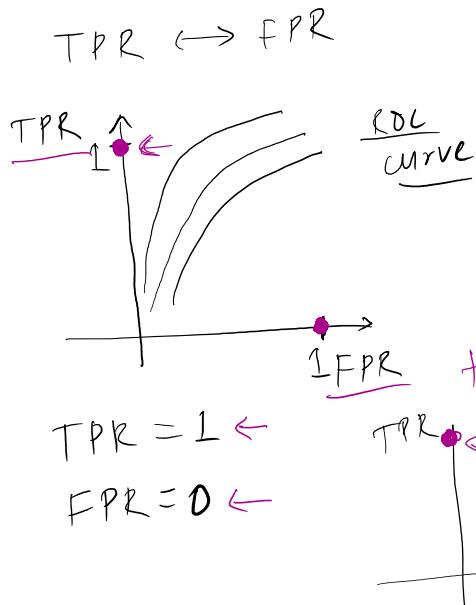
		1	0
1	TP	<u>FN</u>	
	<u>FP</u>	TN	
0			

best case \rightarrow

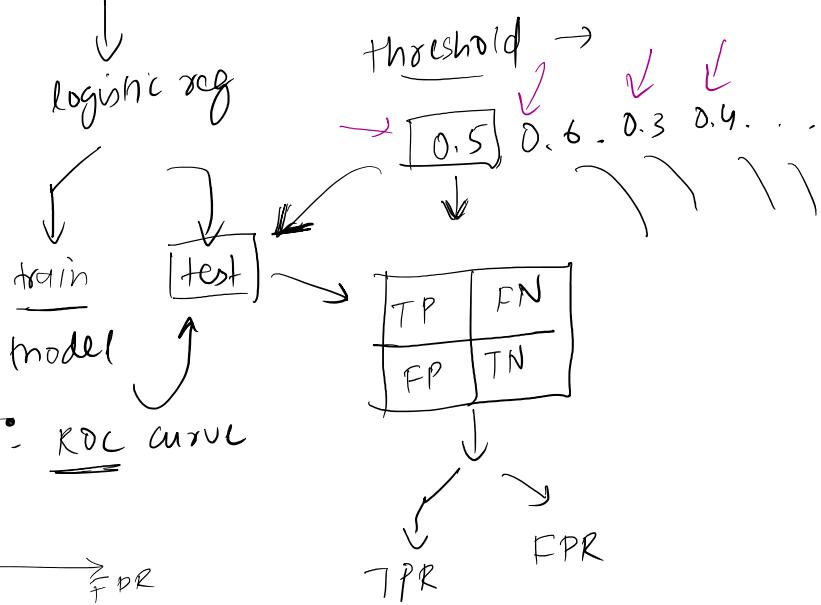
<u>100</u>	<u>0</u>
<u>0</u>	<u>100</u>

ROC Curve

06 June 2023 12:49

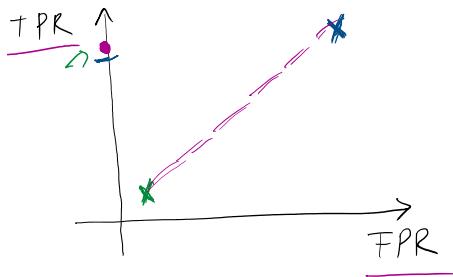


cgpa | iq | placement



Different Cases

06 June 2023 12:49



	1	0
1	TP ↓	FN ↑
0	FP ↓	TN ↑

Predicted

spam $\underline{0.95} >$ spam
email spam

mails \rightarrow spam

$[0.1] \rightarrow$

$[0.2] \rightarrow$ spam

$0.09 \rightarrow$ not spam

$$TPR = \frac{TP}{TP + FN}$$

increase

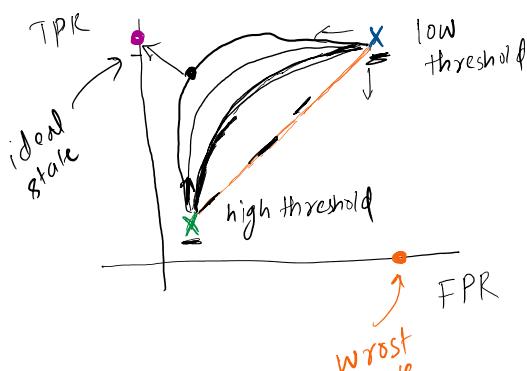
$$FPR = \frac{FP}{FP + TN}$$

increase

threshold → FPR ↑
TPR ↓



decrease threshold → increase threshold



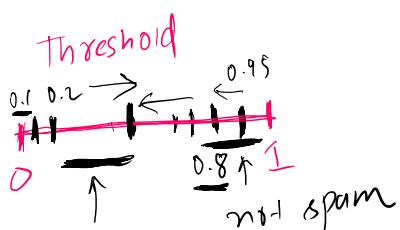
	1	0
1	TP ↑	FN ↓
0	FP ↑	TN

Prediction

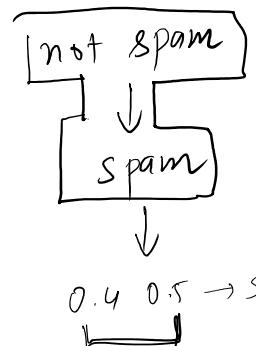
TPR

FPR

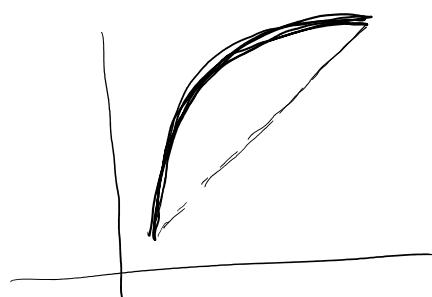
$$[TPR] = \uparrow$$



TPR ↓ FPR ↓ FPR not mu
threshold → $[0.8] \rightarrow$ relax → spam
 $0.95 \rightarrow$ prob - spam
TPR ↑ FPR ↑ $0.8 \rightarrow$ spam



$0.4 \ 0.5 \rightarrow$ spam



Code Example

06 June 2023 12:49

AUC-ROC

06 June 2023 12:49

The AUC-ROC measures the entire two-dimensional area underneath the entire ROC curve from (0,0) to (1,1). AUC provides an aggregate measure of performance across all possible classification thresholds.

- An AUC of 1 indicates that the model has perfect discrimination: it correctly classifies all positive and negative instances.
- An AUC of 0.5 suggests the model has no discrimination ability: it is as good as random guessing.
- An AUC of 0 indicates that the model is perfectly wrong: it classifies all positive instances as negative and all negative instances as positive.

In practice, AUC values usually fall between 0.5 (random) and 1 (perfect), with higher values indicating better classification performance.

AUC = 1

model

AUC → compare

