Why ROC curve?

→ reason1:- threshold selection.
 └→ of classification

→ reason2:- AUC-ROC:- by using
-this we can compare two model
by measuring area under curve
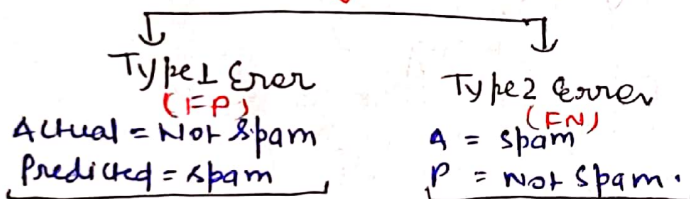(AUC) under ROC

(Normally while giving classification prediction
Model give probability of ~~sign~~ being
(0 or 1))
 ↓
we as Data scientist has to decide
what will be threshold (0.5, 0.7 or 0.6)
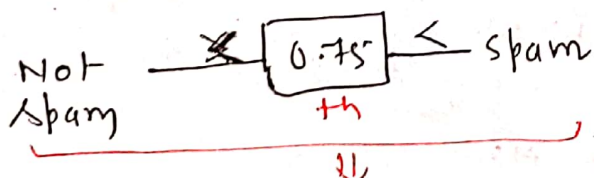 └→ to put probability in target
 class (0, or 1)

Ex

<div style="text-align:center">Model error</div>
 ↓ on threshold = 0.5

Type1 Error	Type2 error
(FP)	(FN)
Actual = Not Spam	A = spam
Predicted = Spam	P = Not Spam.

 ↓
this more
dangerous,
 ↓
to solve this.] ⇒ we will adjust
 our threshold such
 a way that

Not ——✗—[0.75]—←— Spam
Spam   th
              ↓

this way we can
handle our precision &
recall value with our
intuition.

Q. how we will get correct
threshold for our business
problem?

 └→ this answer will be provided
 by [RoC curv]

represents with a
 accha model है।
True positive Rate (TPR) → Benefits ← to remember

 └→ $TPR = \dfrac{TP}{TP + FN} = \dfrac{\text{Predicted positive}}{\text{Total Actual positive}}$

Pred

TP	FN
FP	TN

(Act on left side)

TPR is also known
as benefit i.e. we
made system to get
. benefit i.e. Nothing be TPR

कितना ख़राब model है।
 ↑
False Positive Rate (FPR) → Cost (Loss)
 ↓ to remember

$FPR = \dfrac{FP}{FP + TN} = \dfrac{\text{wrongly Predicted positive}}{\text{Total Actual Negative}}$

Ex:- Out of all "Not Spam" how many
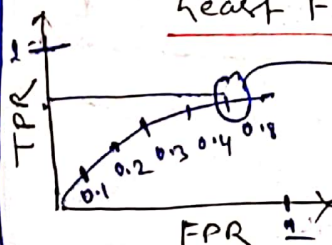 predicted as "Spam"

***

Summary
Best Case
TPR = 100% → when FN = 0
FPR = 0% → when FP = 0

100	0
0	100

ROC curve

 └→ plotting curve b/w TPR & FPR
 for diff-diff threshold to
 decide which threshold is
 ~~best~~ giving max^m TPR &
 least FPR.

→ If this
 threshold is
 giving max^m
 TPR then
 we will select
 this threshold.

Different Cases of threshold
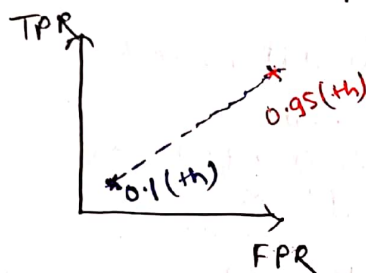
threshold (↓) → $\left.\begin{array}{c}TPR\\FPR\end{array}\right\}$ both (↑)

threshold (↑) → $\left.\begin{array}{c}TPR\\FPR\end{array}\right\}$ both (↓)

⟹ this Conclude that [relation] [between TPR & FPR is] [linear] →[but] ↓

*** Completely wrong Conception

TPR
*0.95(th)
*0.1(th)
FPR

✗ but we dont have such kind of relation actually.

TPR
FPR
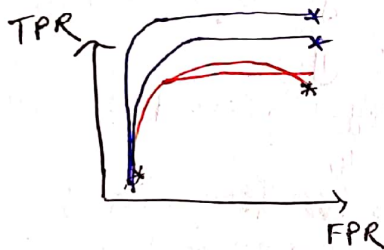
✓ we have this kind of relation ship having TPR & FPR

Practicle Code to bind

finding probability

y_prob = model·predict_Proba(x_test)[:,1]

Finding threshhold, fpr, tpr ?

from sklearn·metrics import roc_curve

fpr, tpr, threshold = roc_Curv(y_test, y_prob)
⟹ st will return FPR & TPR too each threshold.

AUC-ROC

The AUC-ROC measures the entire 2-D area under the entire ROC curv from (0,0) to (1,1).

→ If AUC = 1 ⟹ model is Perfect.

→ If AUC = 0·5 ⟹ model has no discrimination ability. it is as good as random guessing

→ If AUC = 0 ⟹ Model is Perfectly wrong.

∴ AUC usually ranges from [0·5, 1]
∵ 0·5 (random) & 1 (Perfect) with higher values indicating better classification performance.