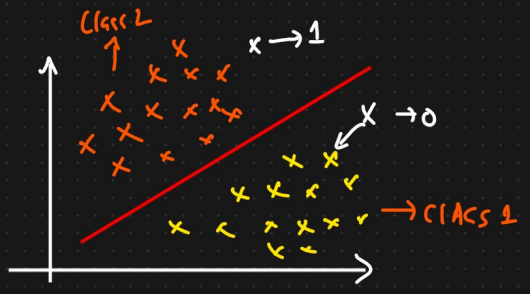
Performance Metrics, Accuracy, Precision, Recall And F-Beta

Topics to be covered

1. Confusion Matrix ✓
2. Accuracy ✓
3. Precision ✓
4. Recall ✓
5. F-Beta Score



Dataset		Actual	Predicted	
x_1	x_2	y	$\hat{y}$	← Model Prediction
–	–	0	1	→ Wrong Prediction
–	–	1	1	→ Correct Prediction
–	–	0	0	→ " "
–	–	1	1	→ " "
–	–	1	1	→ " "
–	–	0	1	→ Wrong Prediction
–	–	1	0	

① Confusion Matrix

y ↑ Actual values

	1	0
1	3	2
0	1	1

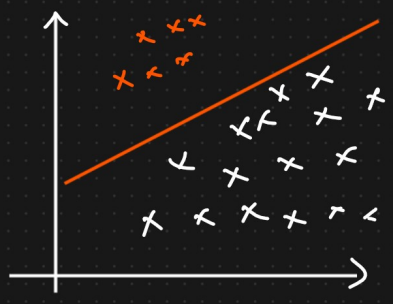
Predicted value → $\hat{y}$ (Actual)

	1	0
1	TP	FP
0	FN	TN

(Predicted)

$$\text{Model Acc.} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$= \frac{3+1}{3+2+1+1} = \frac{4}{7} = 0.571 \approx 57.1\%$$



② DATASET → Imbalanced dataset

1000 datapoints → 900 → 1, 100 → 0 ⇒ Imbalanced dataset

Dumb Model → o/p → 1 ⇒

Accuracy = 90%

Imbalanced Datas...

In this scenario we cannot use Accuracy performance

② Precision = $\frac{TP}{TP+FP}$ } Out of all the actual values how many are correctly predicted

	1	0
1	TP	FP
0	FN	TN

FP → Important

FP ↓↓

③ Recall = $\frac{TP}{TP+FN}$ ⇒ FN ↓↓

⇒ Out of all the predicted values how many are correctly predicted with actual values

	1	0
1	TP	FP
0	FN	TN

Usecase 1 ⇒ **Spam classification**

Mail → (predict) → Spam / Not a Spam

	1	0
1	TP	FP
0	FN	TN

Mail → Spam
Model → Spam
} Good Scenario

FP ↓↓↓

0 ∈ Mail → Not a Spam ↓↓
1 ∈ Model → Spam
} ∈ Blunder

FP is Important

1 ∈ Mail → Spam
0 ∈ Model → Not a Spam
} ⇒ FN

PRECISION PERFORMANCE METRICS.

Usecase 2 ⇒ FN is Important

To predict whether a person has diabetes or not

	Diabetes	No Diabetes → Actual
Diabetes	TP	FP
No Diabetes	FN ↓↓	TN

① Actual → Diabetes
Model → Diabetes
} Good

② Actual → Diabetes
Model → No. Diabetes
} Blunder ⇒ FN ↓↓ ⇒ Important

③ Actual → No Diabetes
Model → Diabetes
} FP ⇒ Wrong Prediction

④ Actual → No Diabetes
Model → No Diabetes
} Correct

⇓

RECALL

Assignment: ① Tomorrow the stock will crash or not

Reduce FP ↓↓ or FN ↓↓

④ F-Beta Score $= (1+\beta^2) \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$

① If FP & FN are both important

$\beta = 1$

Harmonic Mean

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

② If FP is more important than FN

$\beta = 0.5$

$$\text{F 0.5 Score} = (1+0.25) \frac{P*R}{P+R}$$

③ If FN >> FP

$\beta = 2$

$$\text{F2 Score} = (1+4) \frac{P*R}{P+R}$$