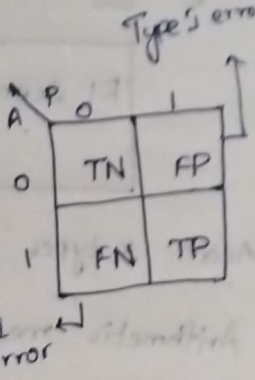
Evaluation Metrics of Logistic Regression:

1. Confusion Matrix:

Binary classification:

0 → -ve (N)

1 → +ve (P)



Negative / Positive will be decided based on Predicted values.

TN, TP → Correctly predicted

FP, FN → Wrongly predicted

FN is actually more dangerous.

FN → Type II error

FP → Type I error

2. Accuracy:

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Overall Predictions}}$$

$$= \frac{TN+TP}{TN+TP+FP+FN}$$

When accuracy?

*Whenever target classes are balanced, we choose accuracy.

3. Recall: (Based on Actual values)

$$\text{Recall } 0 = \frac{TN}{TN+FP}$$

$$\text{Recall } 1 = \frac{TP}{FN+TP}$$

$$\text{Actual Recall} = \frac{\text{Recall } 0 + \text{Recall } 1}{2}$$

4. Precision (Based on Predicted values):

$$\text{Precision } 0 = \frac{TN}{TN+FN}$$

$$\text{Precision } 1 = \frac{TP}{TP+FP}$$

↳ Wrongness of the model

$$\text{Precision} = \frac{\text{Precision } 0 + \text{Precision } 1}{2}$$

5. F1 Score:

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

Mean 3 types

Arithmetic mean $= \frac{a_1 + a_2 + \cdots + a_n}{n}$

Geometric mean $= \sqrt{a_1 \cdot a_2 \cdot a_3 \cdots a_n}$

Harmonic mean $= \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}} = \frac{2ab}{a+b}$

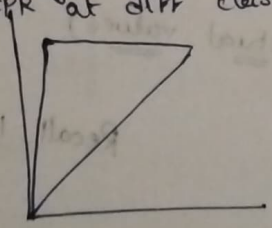
* All classification evaluation metrics (except confusion matrix) ranges from 0 to 1.
* In three mean's harmonic mean gives least value

6. AUC:

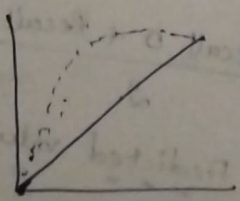
* Area Under the Curve
* AUC always 1 (total)
* AUC Range 0 to 1.

7. ROC: Receivers Operating characteristic.
Graphical representation of TPR & FPR at diff classification thresholds.

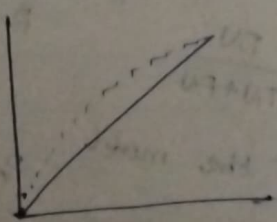
ROC = 1 Perfect



ROC = 0.5 Avg model



ROC < 0.5 worst model



Confusion Matrix for 3 classes:

Actual \ Predicted	H	M	L
H	TH	FM	FL
M	FH	TM	FL
L	FH	FM	TL

$$\text{Accuracy} = \frac{TH + TM + TL}{TH + TM + \dots + TL}$$

$$\text{Recall for M} = \frac{TM}{FH + TM + FL}$$

$$\text{Precision for H} = \frac{TH}{TH + FH + FH}$$