Performance Matrix (Measurement)

1. Accuracy :

$$Acc = \frac{\text{Number of correctly classified pts}}{\text{Total number of Data pts}}$$

Problem with Accuracy :

if the dataset is imbalanced then the accuracy will be affected by majority class.

2. Probability Score :

- Yp will be a probability for any particular class.
- for each class we will calculate the accuracy.
- by keeping some threshold we can make the final prediction.

3. Confusion Matrix :

Confusion Matrix is the visual representation of the Actual VS Predicted values.
It measures the performance of our Machine Learning classification model and looks like a table-like structure.

Actual values

Predicted values		**1**	**0**
	1	**TP**	**FP**
	0	**FN**	**TN**

Elements of Confusion Matrix

It represents the different combinations of Actual VS Predicted values. Let's define them one by one.

TP: True Positive: The values were actually positive and were predicted positive.

FP: False Positive: The values which were actually negative but falsely predicted as positive. Also known as Type I Error.

FN: False Negative: The values which were actually positive but falsely predicted as negative. Also known as Type II Error.

TN: True Negative: The values which were actually negative and were predicted negative

Lets take an example :

Actual Values

Predicted Values		1	0
	1	**540**	**150**
	0	**110**	**200**

In the above matrix, we can analyze the model as :
True positive: 540 records of the stock market crash were predicted correctly by the model.
False-positive: 150 records of not a stock market crash were wrongly predicted as a market crash.
False-negative: 110 records of a market crash were wrongly predicted as not a market crash.
True Negative: 200 records of not a market crash were predicted correctly by the model.

Other Evaluation Metrics associated with it

Accuracy:

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

Recall / Sensitivity:

The recall is the measure to check correctly positive predicted outcomes out of the total number of positive outcomes.

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

Precision:

Precision checks how many outcomes are actually positive outcomes out of the total positively predicted outcomes.

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

F1 score:

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

When to use which metrics for evaluation

Domain-Specific case

→ Stock Market Crash Prediction :

The measure which takes into account this problem is FN and therefore Recall.
So we need to focus on reducing the value of FN and increasing the value of Recall.

→ In most medical cases, such as cancer prediction or any disease prediction we try to reduce the value of FN.

→ Spam Detection :

In this case, we need to focus on reducing the value of FP (i.e when the mail is falsely predicted as spam) and as a result, increasing the value of Precision.

In some cases of imbalanced data problems, both Precision and Recall are important so we consider the F1 score as an evaluation metric.

ROC Curve and AUC score

-- Binary Classification
-- Probability Score

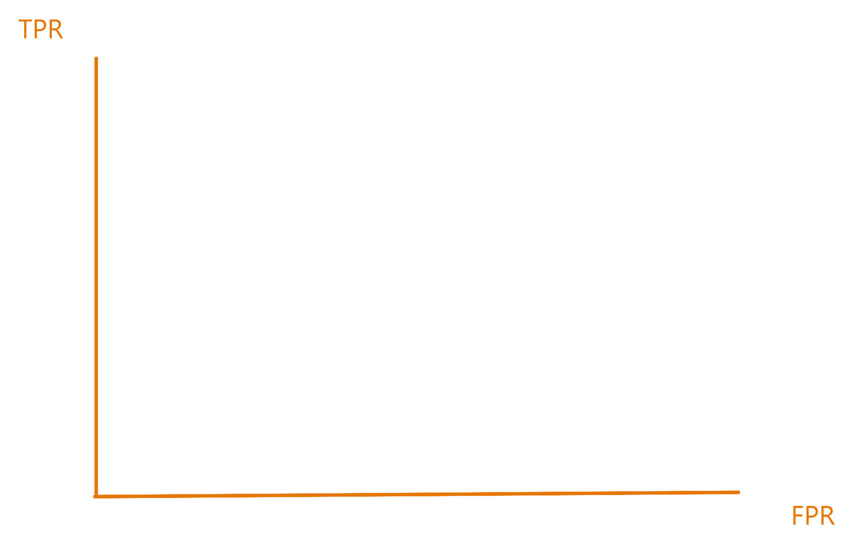
TPR = TP/(TP+FN)

FPR = FP/(FP+TN)

Steps:

1. find the probability score for class.
2. Take each probability as the threshold.
3. calculate TPR and FPR.
4. Plot a graph with TPR and FPR values.

A	B	C	D	E	F	G	H
X1	X2	X3	Y(true)	Y(proba)	t1	t2	t3
1	0	13	1	0.95			
1	1	9	0	0.6			
1	2	16	1	0.75			
0	3	0	0	0.4			
0	14	1	1	0.6			



Key Points :

1. AUC --- 0 → 1
Better

2. if AUC < 0.5 then swap the predicted values.

Log loss

--- it is also based on Probability score (as small as good).
--- For binary classification

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$