

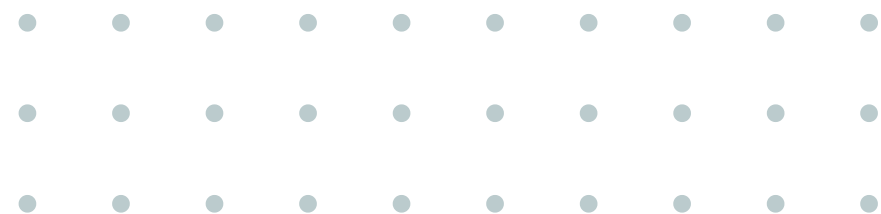
EVALUATION METRICS

CLASSIFICATION ACCURACY

It is the ratio of number of correct predictions to the total number of input samples.

$$\textit{Accuracy} = \frac{\textit{Number of Correct predictions}}{\textit{Total number of predictions made}}$$

It works well only if there are equal number of samples belonging to each class.



CONFUSION MATRIX

It is matrix that describes the complete performance of model

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



There are 4 important terms :

- **True Positives** : The cases in which we predicted YES and the actual output was also YES.
- **True Negatives** : The cases in which we predicted NO and the actual output was NO.
- **False Positives** : The cases in which we predicted YES and the actual output was NO.
- **False Negatives** : The cases in which we predicted NO and the actual output was YES.

Accuracy for the matrix can be calculated by taking average of the values lying across the “main diagonal” i.e

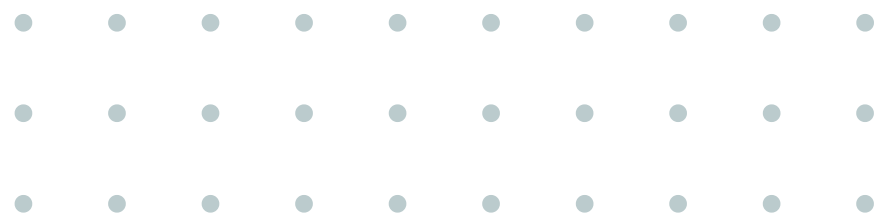
$$Accuracy = \frac{TruePositive + TrueNegative}{TotalSample}$$

F1 SCORE

F1 Score is used to measure a test's accuracy

F1 Score is the Harmonic Mean between **precision** and **recall**. The range for F1 Score is [0, 1].

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

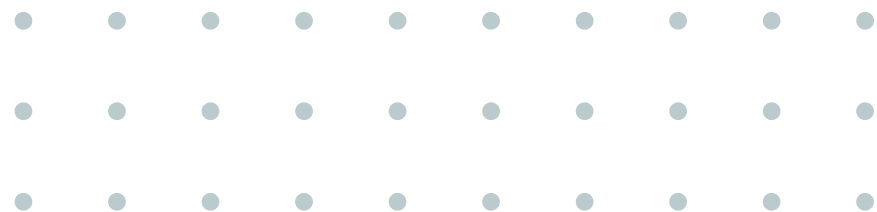




PRECISION

It is the number of correct positive results divided by the number of positive results predicted by the classifier.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

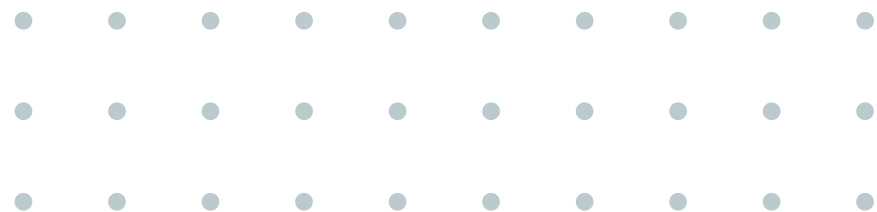


RECALL



It is the number of correct positive results divided by the number of positive results predicted by the classifier.

$$\text{Recall} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}}$$

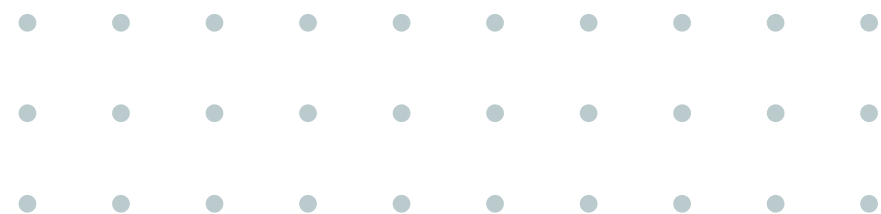


RECEIVER OPERATING CHARACTERISTIC(ROC) CURVE

ROC curve represents a relation between True Positive Rate(TPR) (recall) and False Positive Rate (FPR)(NOT precision).

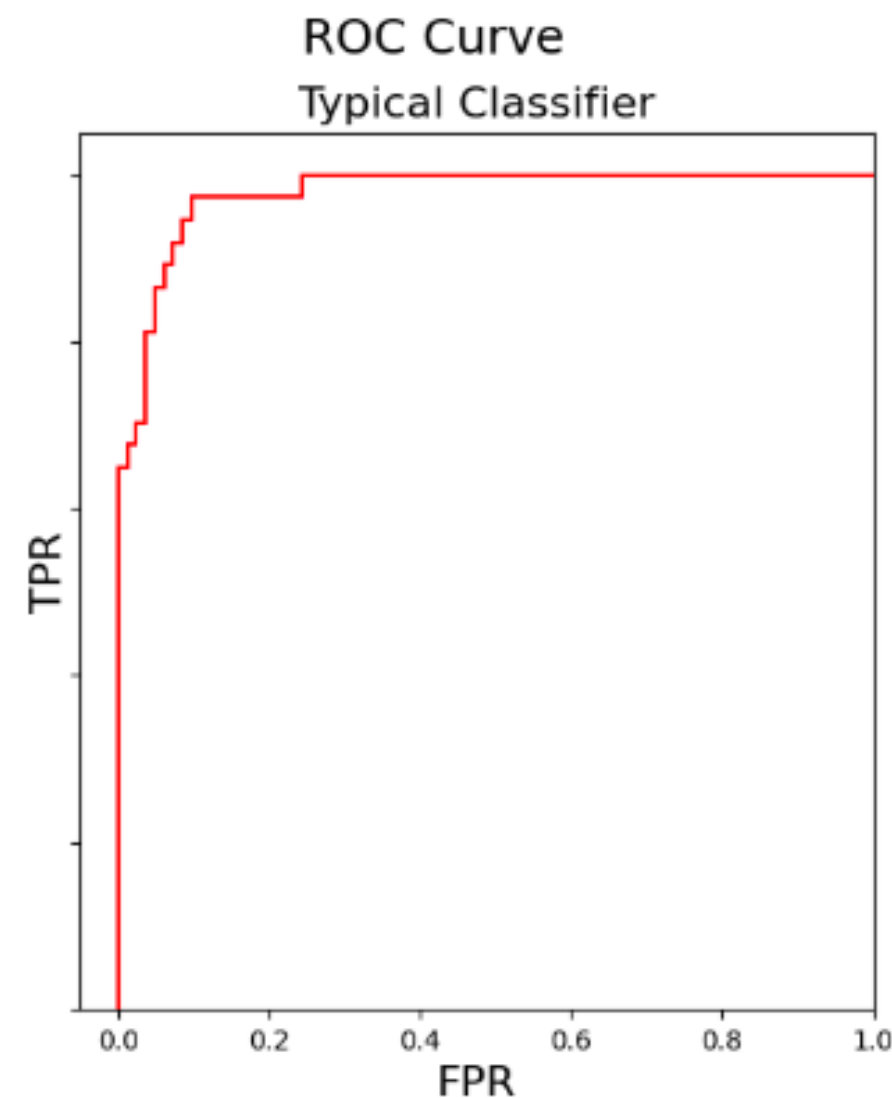
$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

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



The ROC curve is nothing but just the plot between TPR and FPR at different values of threshold.

When we plot the TPR on the y-axis and FPR on the x-axis, we obtain the following ROC curve.



AREA UNDER CURVE(AUC)

It is used for binary classification problem.

Area Under Curve or simply Area Under ROC curve is an alternate representation of the ROC curve in a numerical form and represents the area under the ROC curve.

