PERFORMANCE METRICES:

It is defined on the probability estimates and measures the performance of a classification model where the input is a probability value between '0' and '1'.

→ The performance metrices can be understood more clearly by differentiating it with accuracy.

WHAT ARE THE PERFORMANCE METRICS FOR CLASSIFICATION?

1. Accuracy.
2. Confusion matrix.
3. Precision & Recall.
4. F1 - Score.
5. ROC - AUC

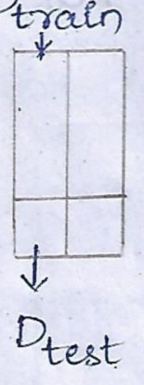
1. ACCURACY:

It is the number of correctly predicted data points out of all the datapoints.

→ For Example,

If the algorithm classified a false data point as true, it would be a false positive.

$$\boxed{\text{Accuracy} = \frac{\text{No. of correctly classified data points}}{\text{Total No. of data points}}}$$



→ D_{test} ⟹ X_test → f → Y_test_pred

↓

we get them in form of an array.

$$\Rightarrow D_{test} \begin{cases} 60 \text{ +ve} \xrightarrow{f} 53 \text{ +ve}, 7 \text{ -ve} \\ 40 \text{ -ve} \xrightarrow{f} 35 \text{ -ve}, 5 \text{ +ve} \end{cases}$$

$53 + 35 = 88$ $7 + 5 = 12$

↳ Ground Truth

So, there are 7-ve's & 5+ves after the Predictions i.e., the errors made by the algorithm.

→ there are 12 Misclassified points

⇒ Accuracy = $\frac{88}{100}$ = 0.88 (or) 88%

→ If

$$\Rightarrow D_{test} \begin{cases} 90+ve \xrightarrow{f} 90+ve,\ 0-ve \\ 40-ve \xrightarrow{f} 0-ve,\ 10+ve \end{cases} \rightarrow \text{NO USE}$$

Even if the accuracy is 90%, there is only one class.

∴ the Data is totally Imbalanced in this case.

2. CONFUSION MATRIX:

It is a technique for summarizing the Performance of a classification algorithm.

→ A kind of table which helps ~~g~~ us to know the performance of the classification model on a set of test data for that the true values are known.

PREDICTED CLASS

FN/FN+TP
↓
TYPE-II ERROR

		POSITIVE	NEGATIVE	
ACTUAL CLASS	+ve	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)	SENSITIVITY $\frac{TP}{TP+FN}$
	-ve	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)	SPECIFICITY $\frac{TN}{TN+FP}$
		PRECISION $\frac{TP}{TP+FP}$	NEGATIVE PREDICTED VALUE TN/TN+FN	ACCURACY $\frac{TP+TN}{TP+TN+FP+FN}$

TYPE-I ERROR
↓
$\frac{FP}{FP+TN}$

→ Sensitivity is also called as True Positive Rate (TPR) or Recall.

→ Specificity is also called as True Negative Rate (TNR)

WHY DO WE NEED CONFUSION MATRIX?

→ They are used to visualize the important predictive analytics like recall, specificity, accuracy & precision.

→ It is useful because they give direct comparisons of values like TP, FP, TN, & FN.

NOTE:

↳ If The Diagonal values of confusion matrix are high, then the predicted values are good.

3. PRECISION & RECALL :

→ Precision quantifies the number of positive class predictions that actually belong to the positive class.

→ Recall quantifies the number of positive class predictions made out of all positive examples in the data set

↳ Precision is also called as "POSITIVE PREDICTED VALUE".

4. F1 - SCORE :

The F1 - Score, also called as F - score, is a measure of a model's accuracy on a data-set.

→ The F1 - score is a way of combining the Precision and Recall of the model.

↳ So, F1-score is defined as the harmonic mean of the precision and recall.

$$\Rightarrow \boxed{\text{F1-SCORE} = 2 * \frac{\text{PRECISION} * \text{RECALL}}{\text{PRECISION} + \text{RECALL}}}$$

5. ROC-AUC CURVE :

→ ROC - RECEIVER OPERATING CHARACTERISTIC CURVE.

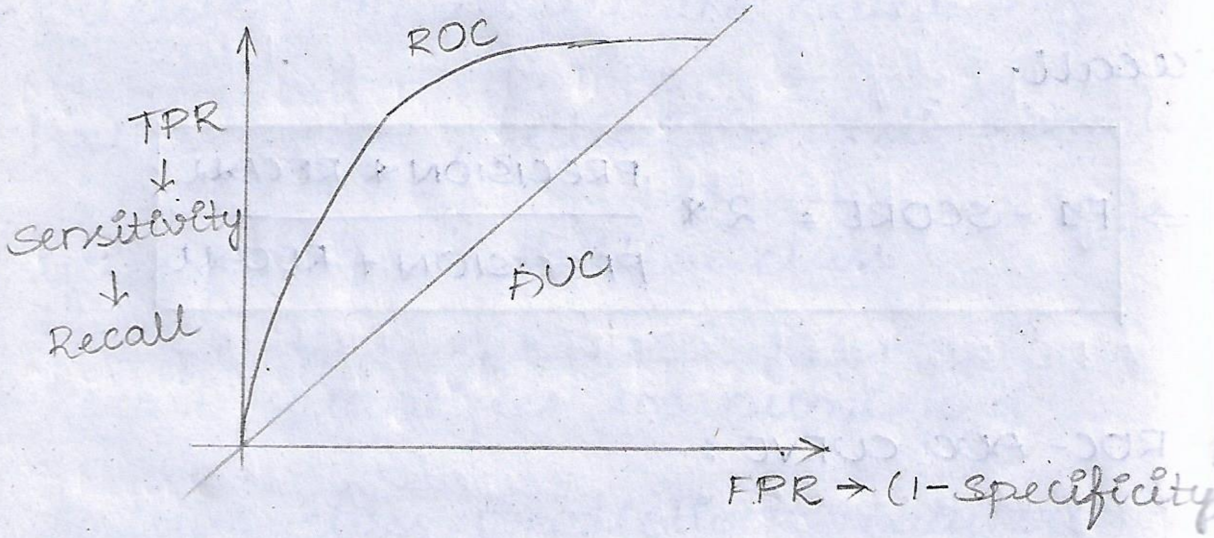
→ AUC - AREA UNDER CURVE.

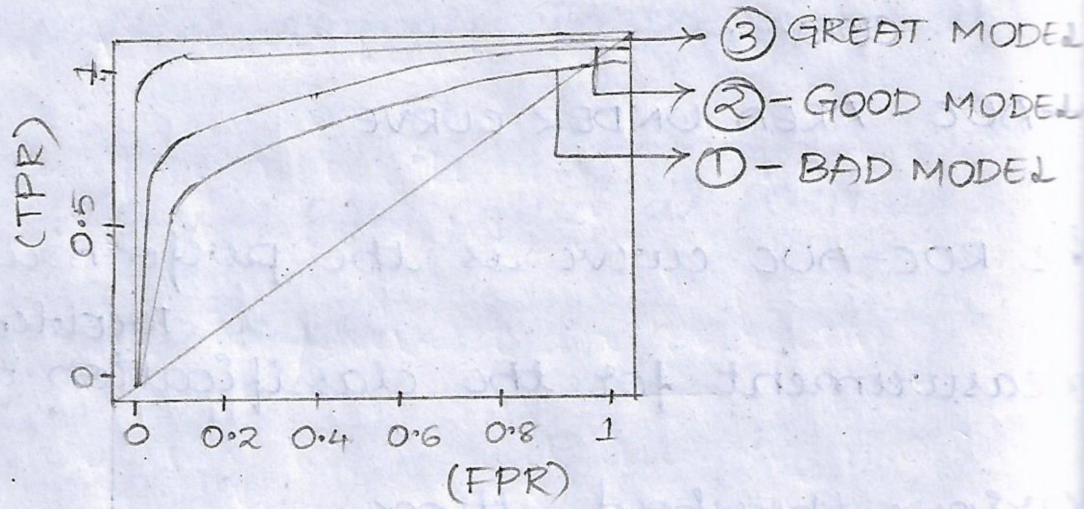
The ROC-AUC curve is the performance measurement for the classification problems at various threshold settings.

→ ROC is a probability curve & AUC represents the degree (or) measure of the seperability.

↳ It tells how much that the model is

capable of distinguishing between the classes.





NOTE:

→ It only works on BINARY CLASSIFICATION

↳ Ideally, we would like to have high sensitivity & high specificity. But in real world scenarios, there is always a trade-off between sensitivity & specificity.

Some important characteristics of ROC-AUC :

→ The value can range from '0' and '1'.

→ However AUC score of a random classifier for balanced data is 0.5

→ ROC-AUC score is independent of the threshold set for classification because it only considers the rank of each prediction and not its absolute value.

→ The same is not true for F1-score which needs a threshold value in case of probabilities output.