Evaluation Metrics for Classification:

Confusion Matrix →

Actual Values		Predicted Values: True	Predicted Values: False
	True	True Positive	False Negative Type 2 Error
	False	False Positive Type 1 Error	True Negative

→ TP + FN = Actual positive values

→ FP + TN = Actual Negative Values

Predicted Positive values = TP + FP

Predicted Negative Values = FN + TN

Using the confusion matrix, we can measure different types of evaluation metrics for classification model.

Accuracy:-

The accuracy of the model is given by the total number of correct predictions divided by the total number of predictions.

It is important to note that in the sklearn module, the positive and negative values of the matrix are reversed; it interprets the classification matrix in this order:

Actual	Negative	Positive
Negative	TN	FP
Positive	FN	TP

If we want to use the original confusion matrix inside sklearn, then there is a label argument inside the confusion matrix function, we set its value to [1, 0]

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

Misclassification error:-

Misclassification error is given by 1 - Accuracy.

When we do model building in real life, the majority of cases the data is imbalanced. If we just use the accuracy to measure performance of the model, then in that case, the accuracy is not a correct measure of the performance of the model, because the data is imbalanced and biased. Here we cannot rely on the accuracy, hence it is called as accuracy paradox.

True Positive Rate (TPR):-

True positive rate tells us that out of the actual positive values, how many were predicted to be by the model. TPR is given by,

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

False Negative Rate (FNR):-

Similar to the True Positive Rate, we also have false negative rate. It is the ratio of the the values which were actually positive but which were predicted negative by the model to the total actual positive values.

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

N.B

High value of TP minimises the FN values. It is better to have True Positive Rate as high as possible and FNR as low as possible to improve the model performance.

True Negative Rate (TNR):

Similar to True positive Rate, True Negative Rate tells us how many actual negative values been correctly predicted by our model. It is given by:

$$TNR = \frac{TN}{TN + FP}$$

False Positive Rate (FPR):

False Positive Rate is given by the ratio of the values which were negative but the model predicted them positive to the total actual negative values.

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

Again similar to The true positive rate, we want true negative rate should be as high as possible.

According to different problem statements and real life scenarios, we decide that what are the metrics that we want to minimise or what are the metrics that we want to increase for the better prediction of our model.

True Positive Rate is also called as Sensitivity.

True Negative Rate is also called as Specifity.

Precision:-

Precision is given by the ratio of actual positive to predicted positive.

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

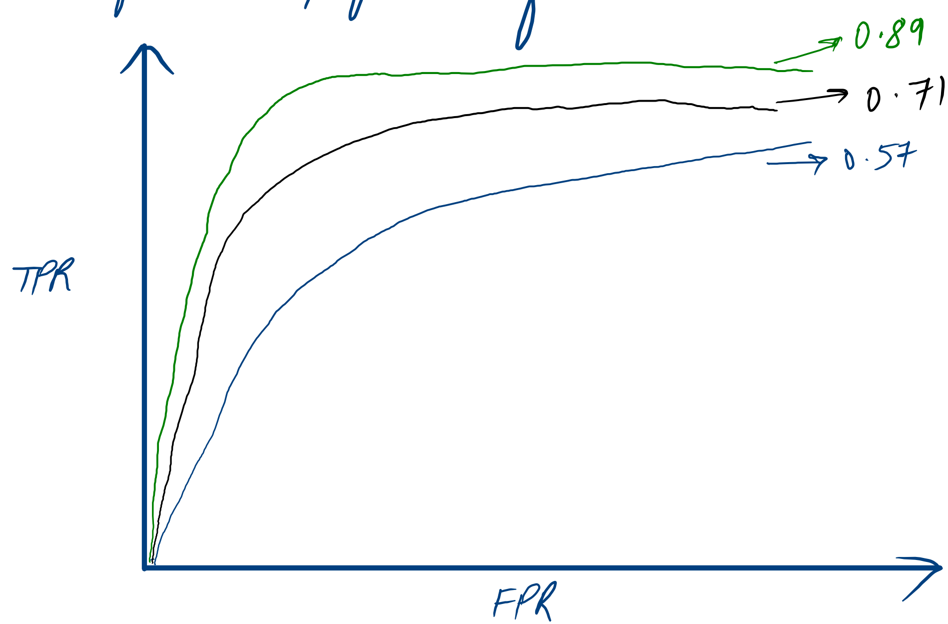
Recall:-

True positive Rate or Sensitivity are also known as Recall.

The choice that what we want to increase and what to decrease completely depends upon the scenario and problem statement.

AUC - ROC (Receiver Operating Characteristics)

In the ROC curve, we plot the FPR and TPR. FPR is plotted on the x-axis and TPR is plotted on the y-axis. AUC is the area under the curve. The area under the curve should be as close to 1 as possible for a better performance of the model.



Now similar to the accuracy score for a classification model, we also have the F1-score.

F-1 Score :-

F-1 score for a classification model is a measure of its performance. Accuracy score gives us just a simple measure that out of true positive values how many were predicted correctly. It does not take into account the imbalance of data or the problem statement at hand. On the other hand, F-1 score is a function or calculated using the precision and recall. It takes into account the imbalance in data and also considers that which metric is important according to the problem statement.

F1-score is given by →

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