Getting Started with Machine Learning

Raunak J

In last tutorial we learned Confusion Matrix. (Pre-requisite)

In this tutorial we'll learn Precision and Recall

Precision :

Fraction of relevant instances among the retrieved instances.

Also we can say the measure of items are relevant out of selected.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall :

Fraction of relevant instances retrieved.

Also we can say the measure of relevant items selected.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Although this works for all classifiers, we'll consider our Logistic Regression Example :

Number of hours dedicated and outcome one gets in pass/fail manner of examination.

$y = [0,0,0,1,1,1]$

Predicted Probabilities after Sigmoid we got were :
$y\_pred = [0.47, 0.53, 0.57, 0.70, 0.64, 0.76]$
After applying threshold :
$y\_pred = [0,1,1,1,1,1]$

Then we calculate Confusion Matrix :
True Positives = 3
True Negatives = 1
False Positives = 2
False Negatives = 0

Now let us calculate Precision and Recall :

$$Precision = \frac{TP}{TP + FP} = \frac{3}{3 + 2} = \frac{3}{5} = 0.6$$

$$Recall = \frac{TP}{TP + FN} = \frac{3}{3 + 0} = \frac{3}{3} = 1$$