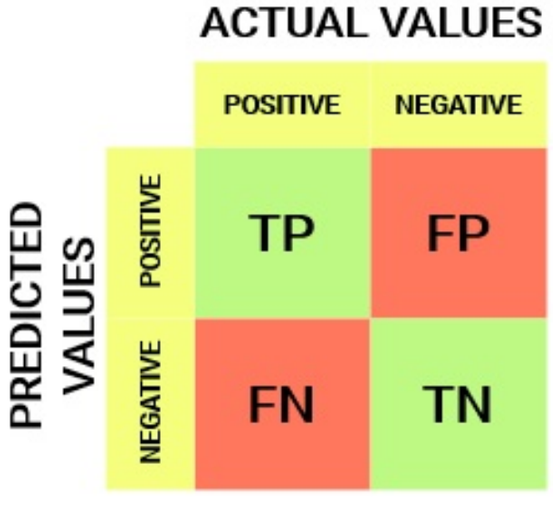
Evaluation Matrix

Confusion Matrix [Evaluation matrix used for classification Problem]

n X n matrix
; n = number of classes in the variable



Type-1 error
FPR = FP / FP+TP

Focus should be to reduce type-1 & type-2 error

Type-2 error
FNR = FN / TP+FN

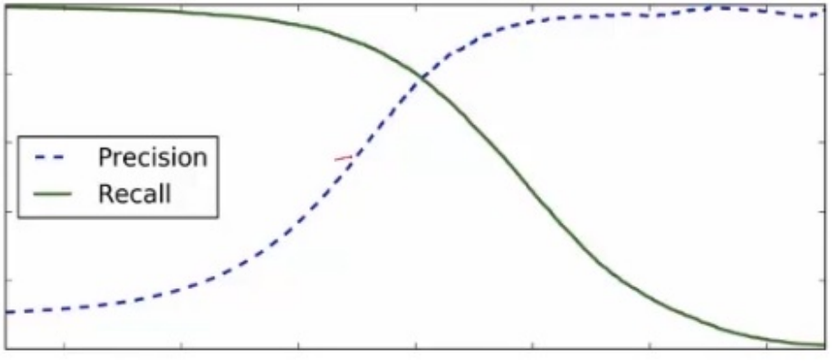
Accuracy = TP + TN / (TP + FN + FP + TN)
(for balanced dataset)

Recall = TP / TP + FN
(Sensitivity)
Precision = TP / TP + FP
(+ve pred value)

for Imbalanced dataset

F1 score

- High Precision, Low Recall
- High Recall, Low Precision
- Choice depends upon the use case
- Combined using F1 Score
- F1 is maximum when precision = recall



$$F_1 = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}$$

What if we have a probabilities of a class (i.e logistic regression) rather than the value for it(0 or 1)

for this sort of problem

we can set threshold value and then decide the classes based on that

ID	Actual Values	Predicted probabilities	At threshold 0.6	
ID1	1	0.9	1	TP
ID2	0	0.51	0	TN
ID3	1	0.47	**0**	**FN**
ID4	1	0.32	**0**	**FN**
ID5	0	0.1	0	TN
ID6	1	0.94	1	TP
ID7	1	0.78	1	TP
ID8	0	0.56	0	TN

"Thresholding"

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$$