

ROC Example:

We have This simple example of ten observations and we also have a classifier, it could be any classifier. This classifier gives us what is the probability of the observation belong to class 1 $P(1|A)$ in Table 1 the observations are sorted from highest probability to lowest. **True class** is the actual class, which is from training set.

We are going to apply this rule:

If $P(1|A)$ is greater than t then $\hat{Y} = 1$ else $\hat{Y} = 0$

t is a threshold.

Table 1: Classifier probabilities and target variable values in training dataset

Instance	$P(1 A)$	True class
1	0.95	1
2	0.93	1
3	0.87	0
4	0.84	1
5	0.84	0
6	0.84	0
7	0.75	0
8	0.52	1
9	0.44	0
10	0.26	1

Table 2: Threshold for all the possible trade-offs

Class	+	-	+	-	-	-	+	-	+	+	
t	0.26	0.44	0.52	0.75	0.84	0.84	0.84	0.87	0.93	0.95	1.00
TP	5	4	4	3	3	3	3	2	2	1	0
FP	5	5	4	4	3	2	1	1	0	0	0
TN	0	0	1	1	2	3	4	4	5	5	5
FN	0	1	1	2	2	2	2	3	3	4	5
TPR	1	0.8	0.8	0.6	0.6	0.6	0.6	0.4	0.4	0.2	0
FPR	1	1	0.8	0.8	0.8	0.4	0.2	0.2	0	0	0

This table will generate Figure 1.

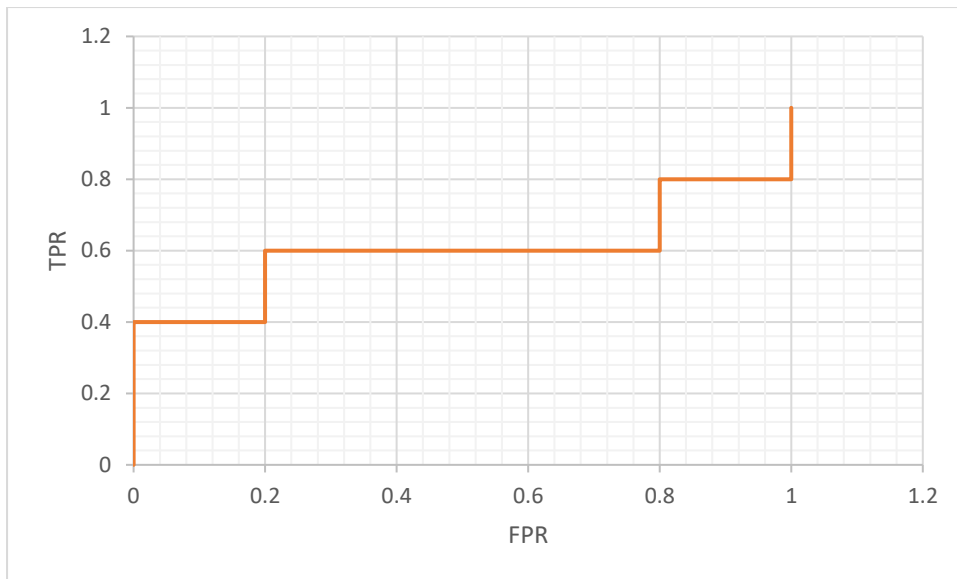


Figure 1: ROC curve

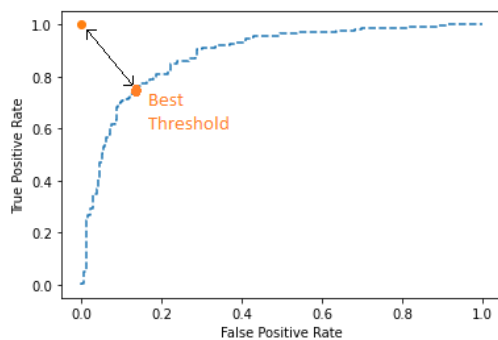
This is how ROC curve could be construct.

Question:

What is the best threshold for this example?

As stated in 'Analytics India Magazine', you want the best threshold to be the closest to 'Perfect Model Point' or (0,1).

Finding optimal threshold from ROC Curve



Source: analyticsindiamag.com

Therefore the **best threshold would be at (0.2,0.6)**, FPR and TPR respectively.