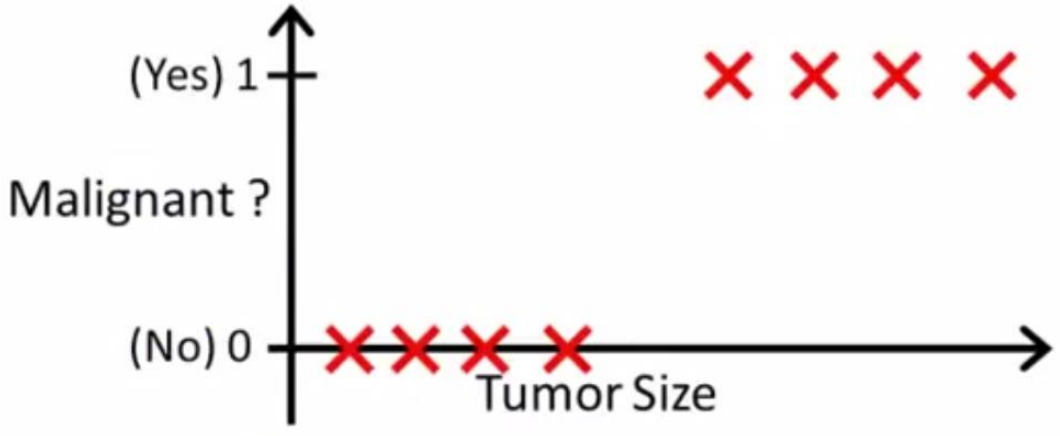
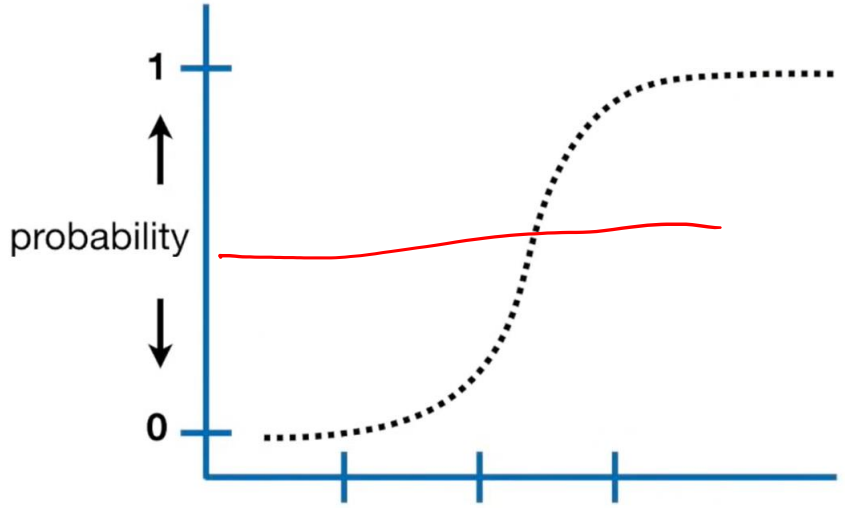
ROC Curve and AUC Score

Dr. Muhammad Wasim

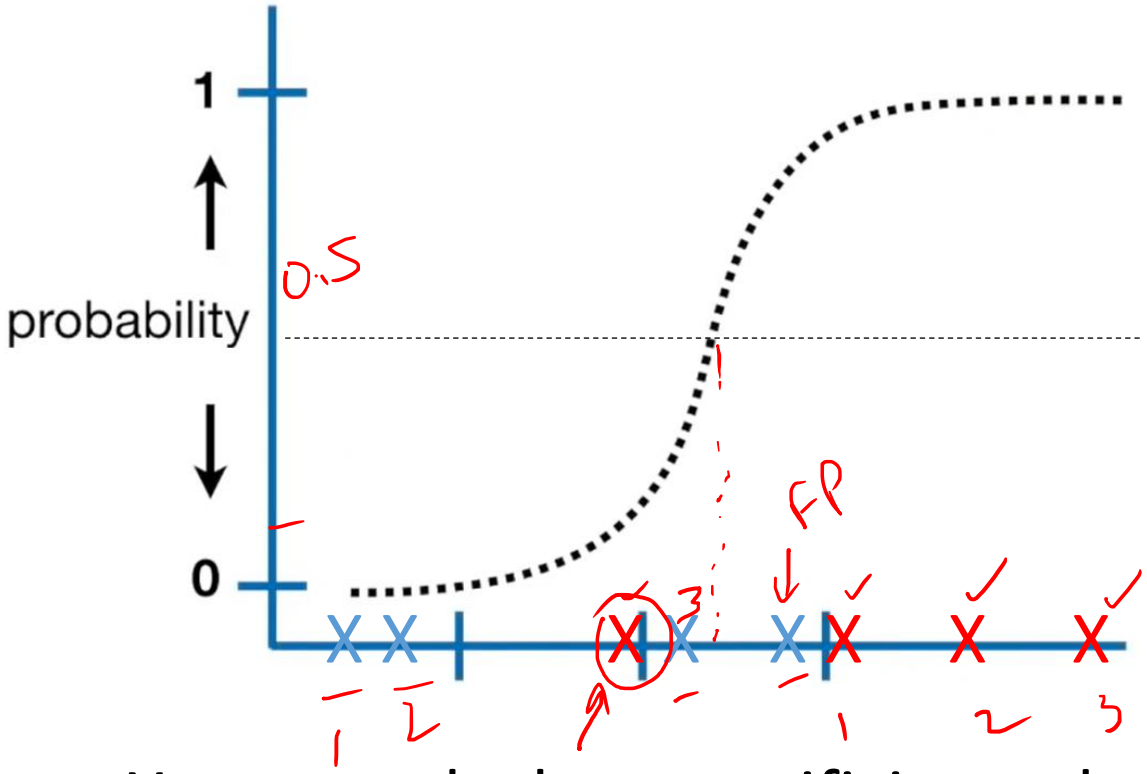
Thresholding for Classification





- How to determine the class of new examples.

Thresholding for Classification



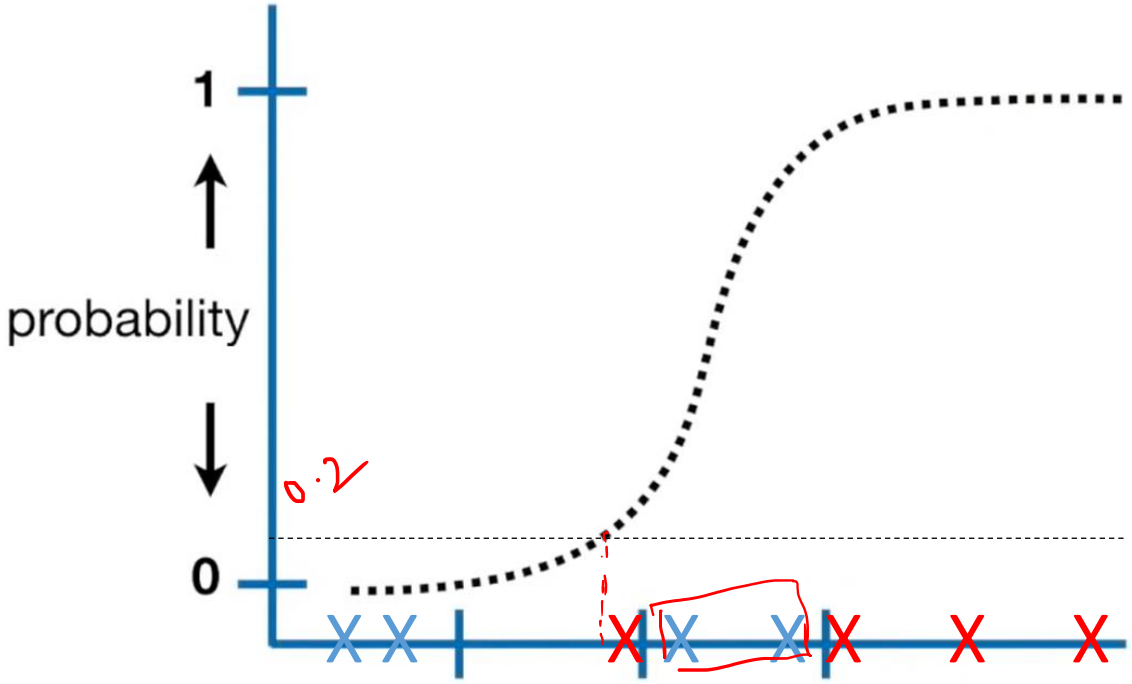
Predicted

Actual

	Malignant	Benign
Malignant	3	1
Benign	1	3

- You can calculate specificity and sensitivity for the threshold set to 0.5

Thresholding for Classification



Actual

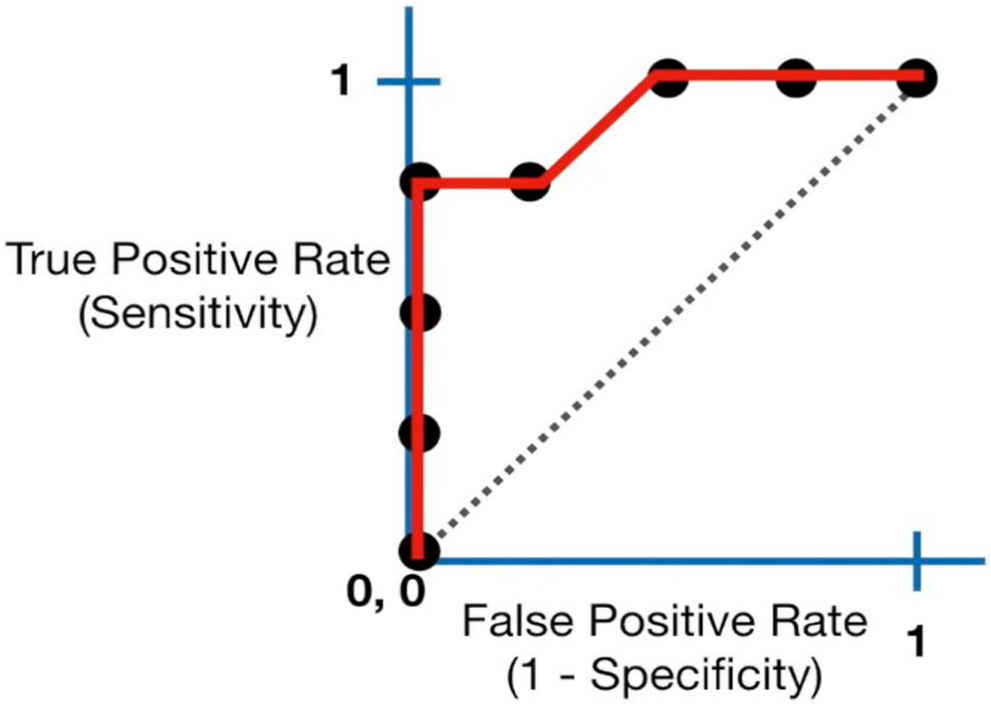
Predicted	Malignant	Benign
Malignant	4	2
Benign	0	2

- It will now classify all patients with cancer correctly. But will also increase false positives.

Thresholding for Classification

- For different thresholds, we need to create confusion matrix for each one.
- It gets difficult to compare the performance of different classification algorithms when we have so many confusion matrices.
- So, instead, we use receiver operator characteristic (ROC) curve which provides a simple way to summarize all information.

ROC Curve



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

$$FPR = 1 - specificity = \frac{FP}{FP + TN}$$

Area Under Curve (AUC)