Classification

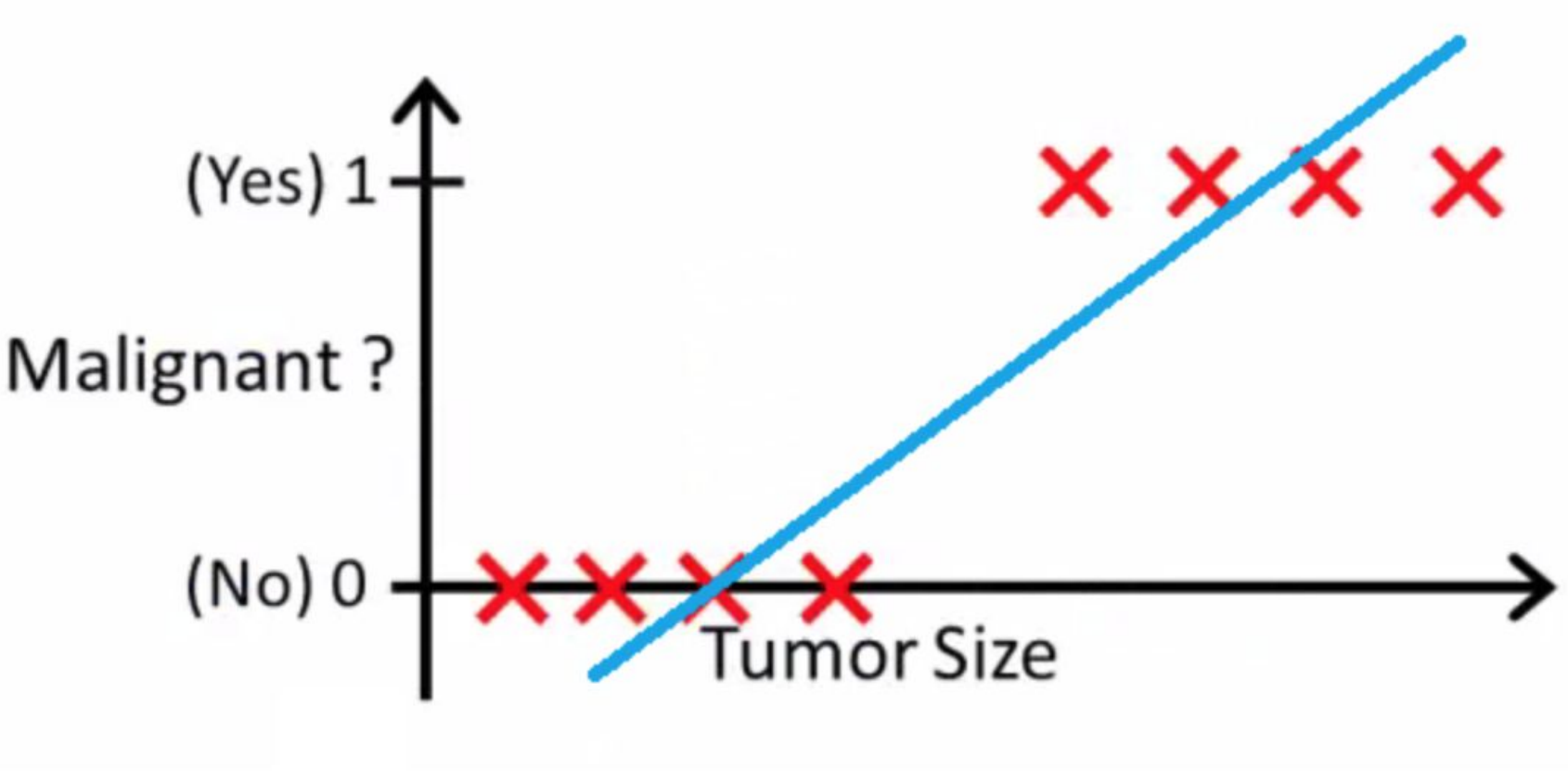
By Saurav Poudel

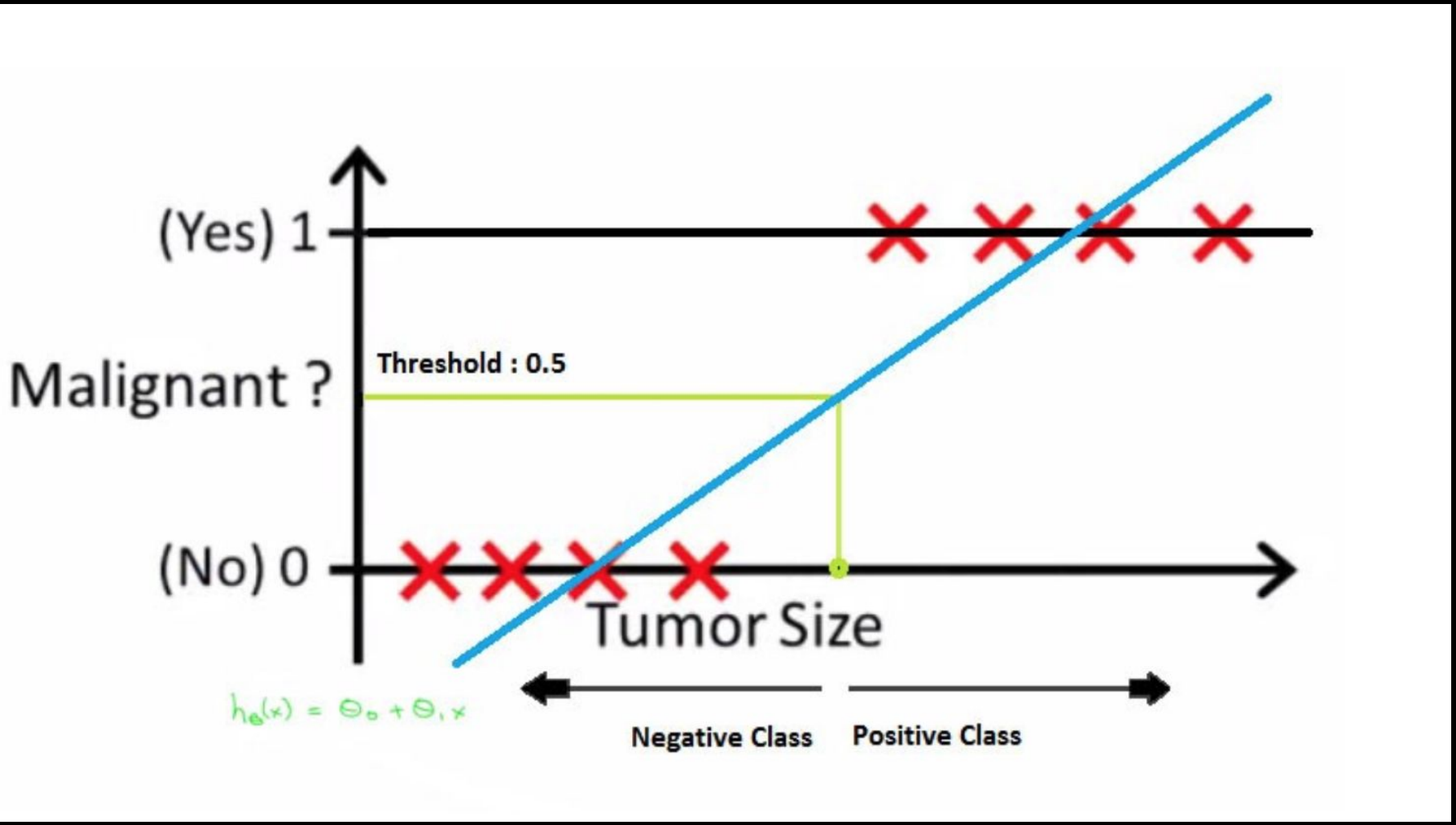
Classification vs Linear Regression

Remember, in LR, we were predicting something that was continuous.

	Two Class Classification	
$y \in \{0, 1\}$	**1 or Positive Class**	**0 or Negative Class**
Email	Spam	Not Spam
Tumor	Malignant	Benign
Transaction	Fraudulent	Not Fraudulent

Why not use the linear regression method?


(Yes) 1
Malignant ?
(No) 0
Tumor Size


(Yes) 1
Malignant ?
Threshold : 0.5
(No) 0
Tumor Size
$h_\theta(x) = \theta_0 + \theta_1 x$
Negative Class
Positive Class


(Yes) 1
Malignant ?
(No) 0
Tumor Size
Negative Class
Positive Class

Probability of Output being 1 given the input.

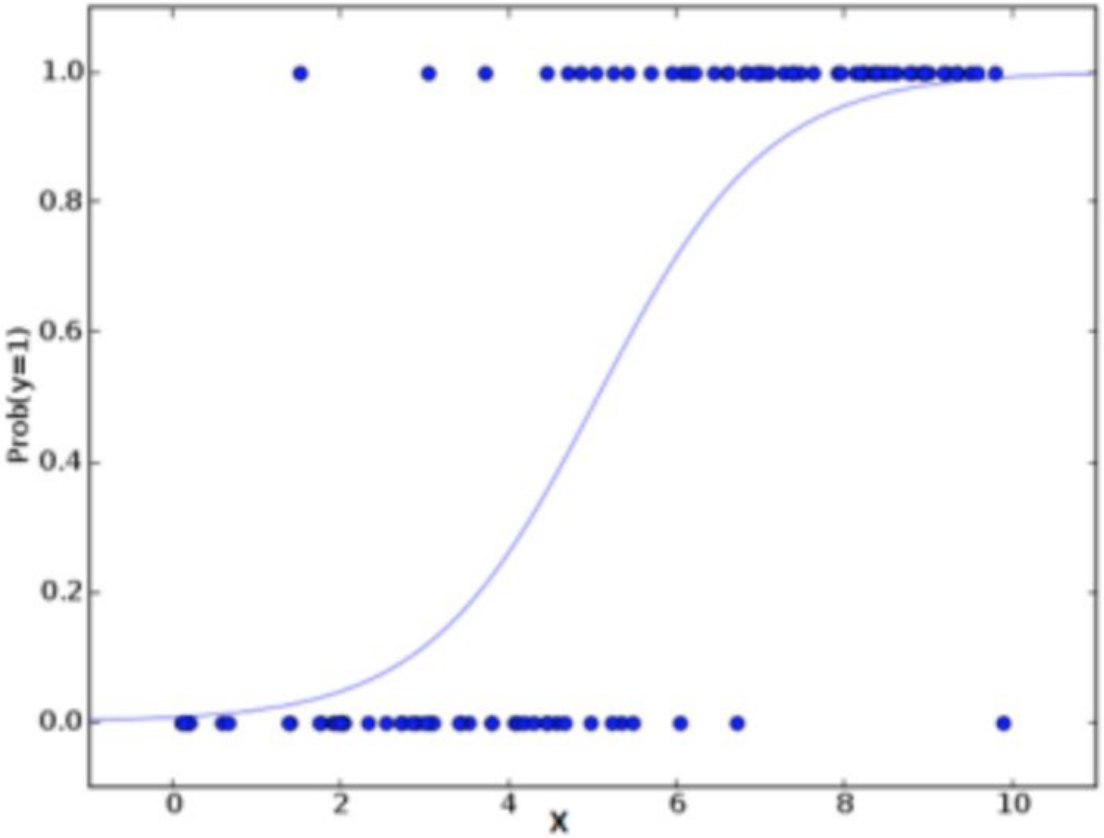
$P(Y=1 \mid X)$

Probability should always be between 0 and 1. Not more or less.

$$p(X) = \beta_0 + \beta_1 X.$$

Sigmoid Function

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}}$$


1.0
0.8
0.6
0.4
0.2
0.0
Prob(y=1)
0
2
4
6
8
10
X

$$p(X) = \beta_0 + \beta_1 X.$$

$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}.$$

Concept of 'Activation Function'

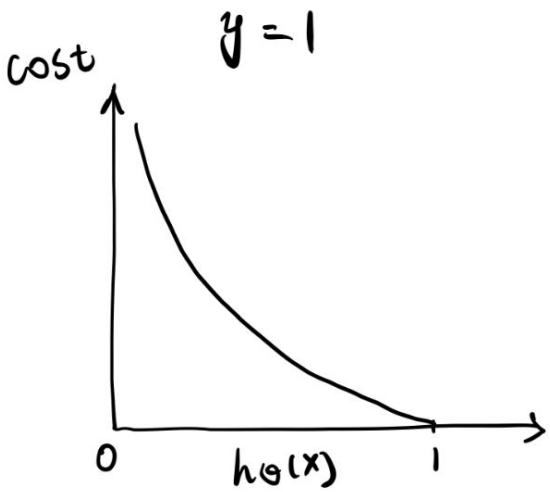
$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X.$$

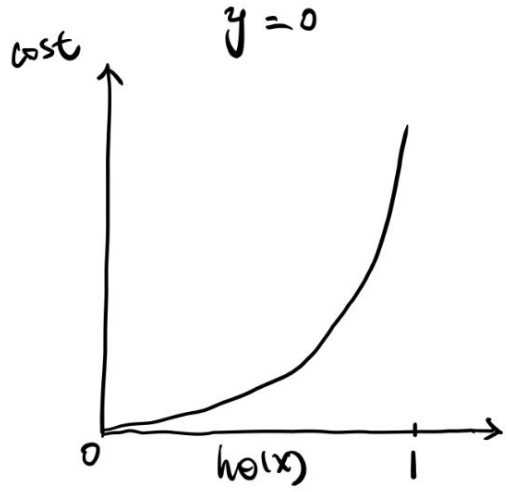
Hence known as "Logistic Regression."

(with the LHS known as Log Odds)

What is the loss function of Logistic Regression?

$$Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{if } y = 1 \\ -log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$





Why didn't we use the MSE as our Loss Function in Logistic Regression?

$P(Y = 1 | x)$ gives a probability only.

So what should be the decision boundary?

How to measure the model accuracy?

Is an accurate model necessarily a good model?

Hypothetical Diabetes Example

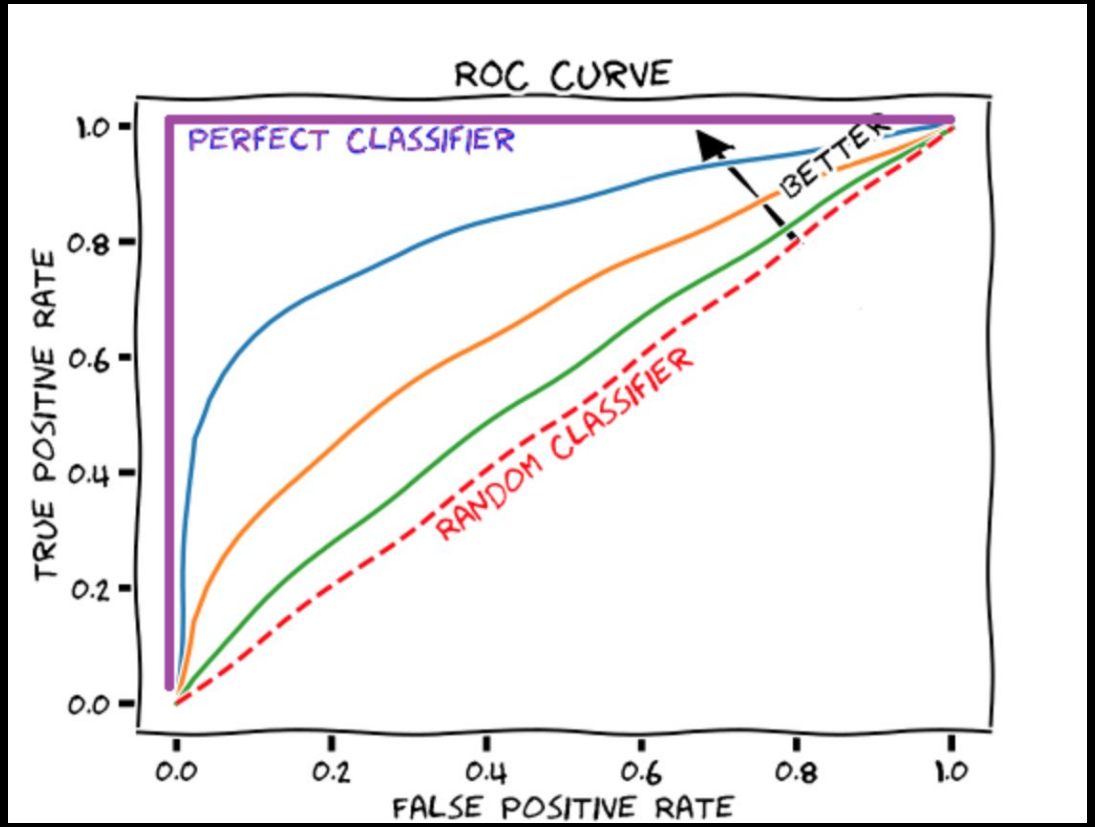
Confusion Matrix

Truth table (confusion matrix)

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

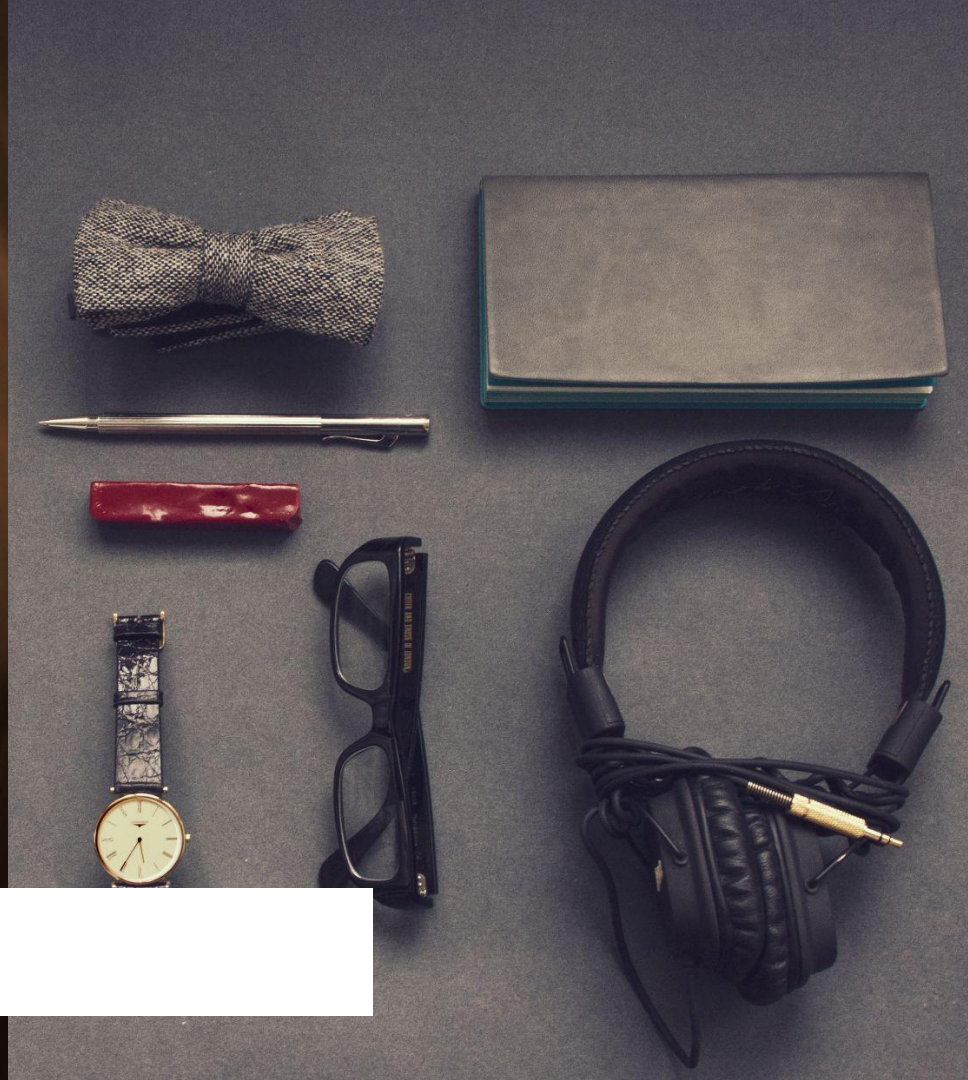
- Accuracy $= \frac{TP+TN}{P+N}$
- Sensitivity $= \frac{TP}{\text{Actual Positives}}$
- Specificity $= \frac{TN}{\text{Actual Negatives}}$

ROC and AUC


ROC CURVE
PERFECT CLASSIFIER
BETTER
RANDOM CLASSIFIER
TRUE POSITIVE RATE
FALSE POSITIVE RATE
0.0
0.2
0.4
0.6
0.8
1.0

Recall

Precision



Let's get into coding now!