LOgisTIC

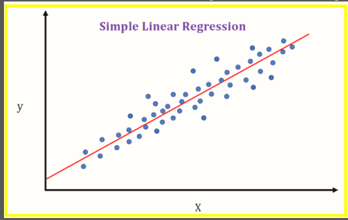
Mi  Regression

Discussion class<u>**C**</u>fication

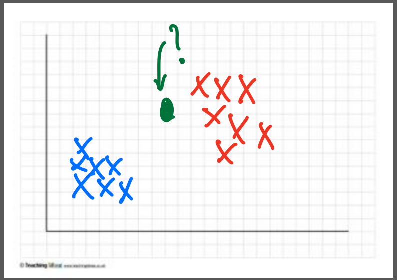
What is logictic regression?

But before that, let's make sure you understand these two terminologies

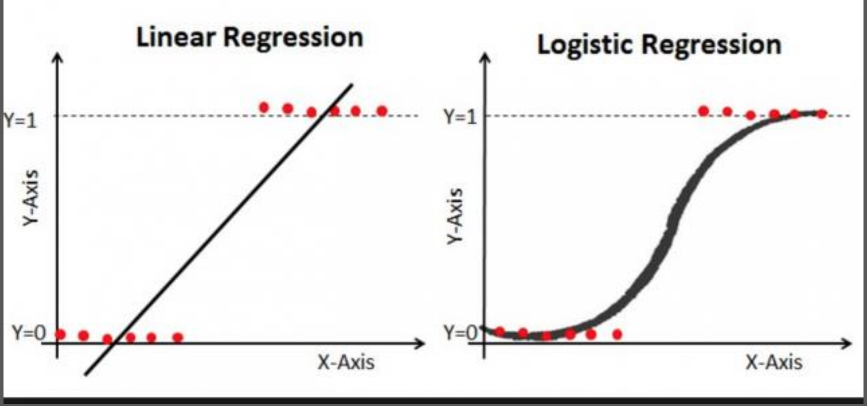
Regression: predicting a continuous quantity output. Basically, finding a trend line

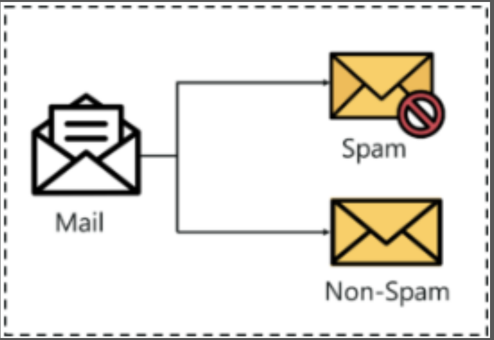


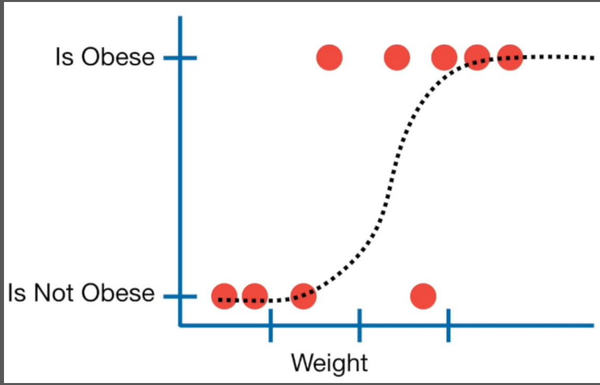
Classification: predicting whether a data belongs to a certain class or not. Basically, find if the output is either in class 0 or class 1

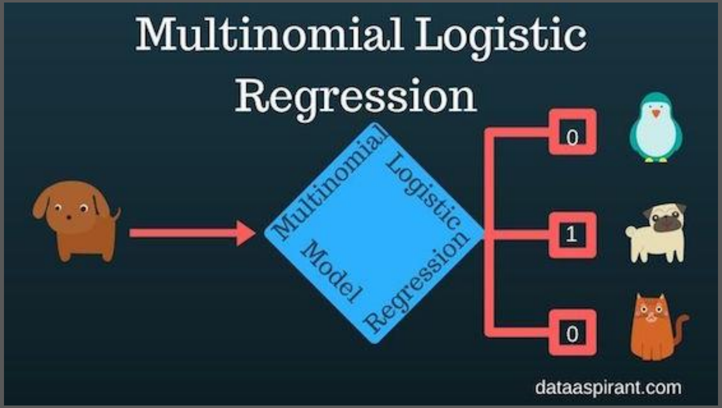


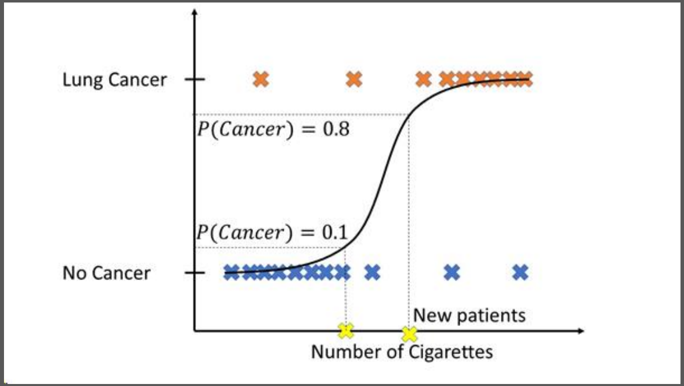
Logistic Regice  ssion




Mail
Spam
Non-Spam


Is Obese
Is Not Obese
Weight

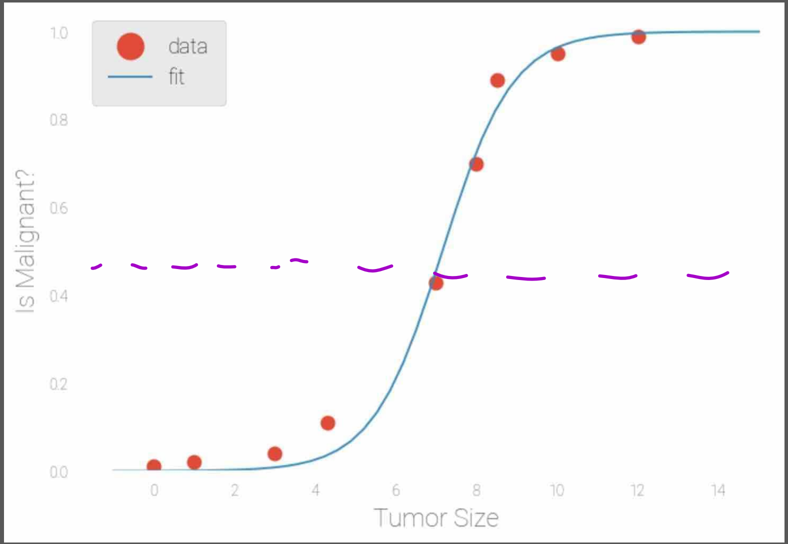

Multinomial Logistic Regression
Multinomial Logistic Regression Model
0
1
0
dataaspirant.com


Lung Cancer
$P(Cancer) = 0.8$
$P(Cancer) = 0.1$
No Cancer
New patients
Number of Cigarettes

Key idea: given input, which class does this input belong to

Elaboration: given input, what is the probability of this input belongs to class A?

Output range is between [0,1]



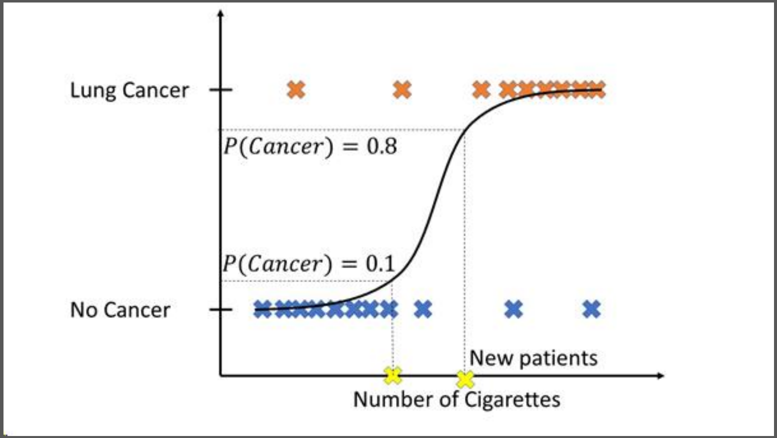
How to determine whether the tumor is malignant?

if $y \geq 0.5 \rightarrow$ malignant

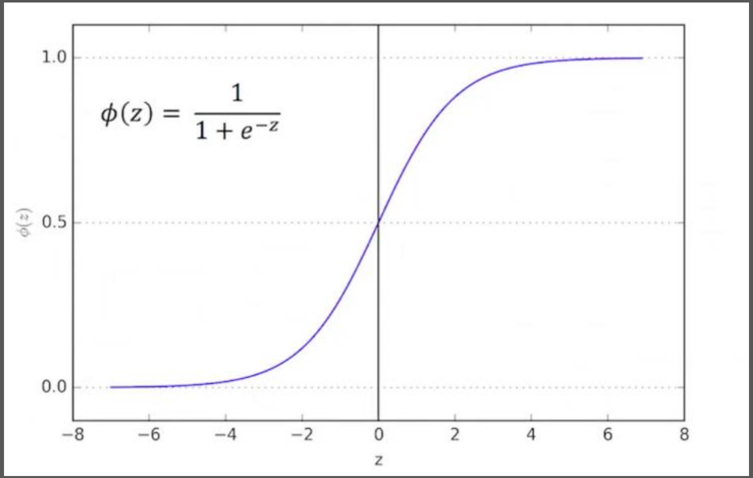
else $y < 0.5 \rightarrow$ benign

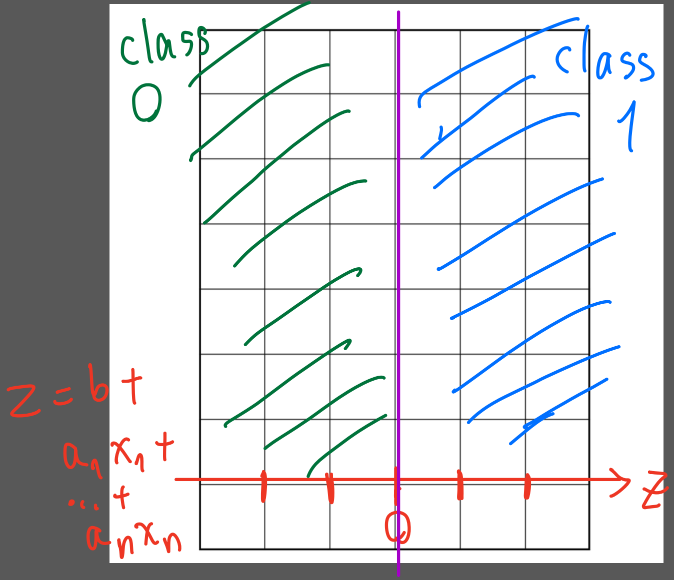
// we can optimize threshold to be something other than 0.5

Problem: see the "jump gap" here? How can we make the curve?



Solution: logistic function





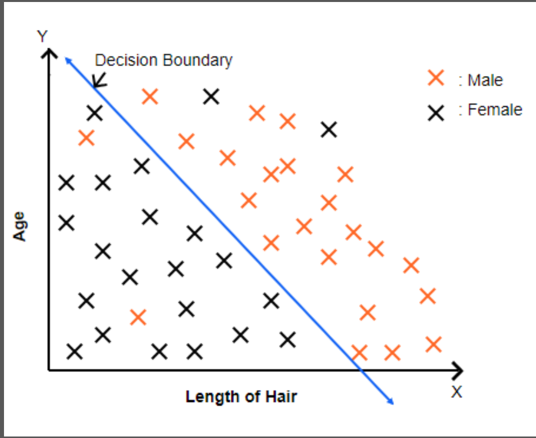
So from this, can we classify which class the input belongs to?

if $z \geq 0 \rightarrow \phi(z) \geq 0.5 \rightarrow$ belong to class 1

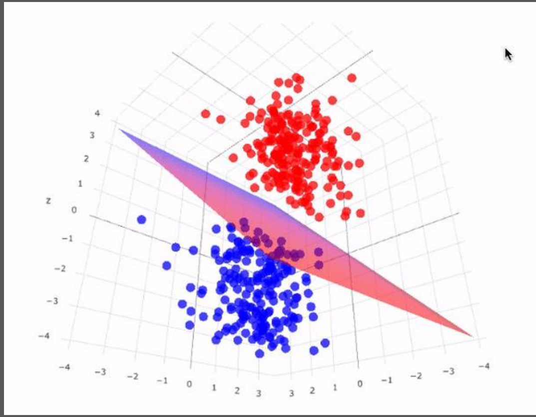
else $z < 0 \rightarrow \phi(z) < 0 \rightarrow$ belong to class 0

sion boundary

Deoxys



↖ 2 variables



↖ 3 variables

Problem: how to we know which threshold is the best one to classify data?

Solution: Likelihoot (likelihood)

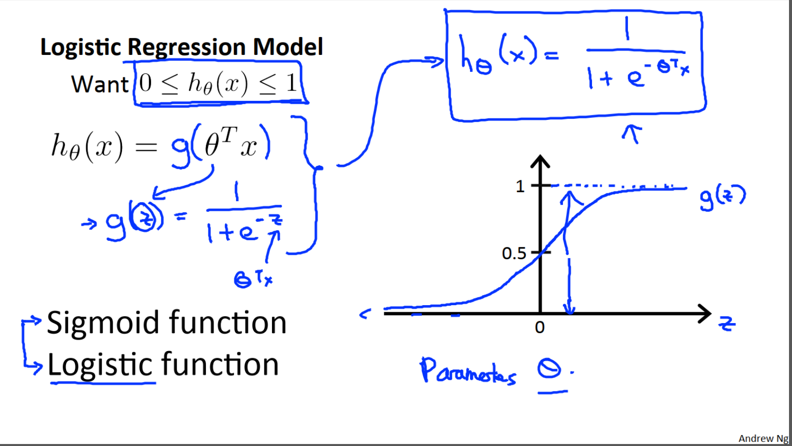
hoothoot

$$L = [\text{product of } \phi(z_i) \text{ of points in class } 1] \times [\text{product of } (1-\phi(z_i)) \text{ of points in class } 0]$$

Goal: We want to maximize likelihood

Note: technically, we want to update weight (theta), and we let z = theta dot x

And our x-axis for sigmoid function is Z

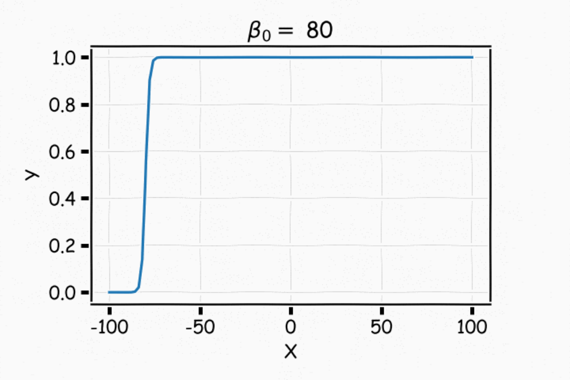


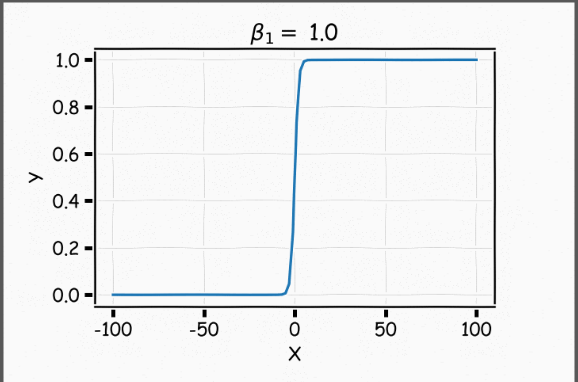
After getting likelihood → update θ → $g_{new}(\theta_{new}^T \cdot x)$

Coefficients of the Logistic Function

$$P(Y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}}$$

- B0 is the intercept
- B1 is the change in log odds (kind of like slope)





(stole this page from stephanie's slide)

- **In general**: log likelihood is another alternative for finding how well the model performs. It is better than likelihood when we want to find the joint likelihood for each data, and we can simplify taking derivative
(just need to know it's easier to do complex math when working with a log)

- To simply put, likelihood is equivalent to logistic regression's cost function

But you know, I learned something today

- Logistic regression is a way to categorize data into a class
- We maximize the likelihood to get the best logistic regression model