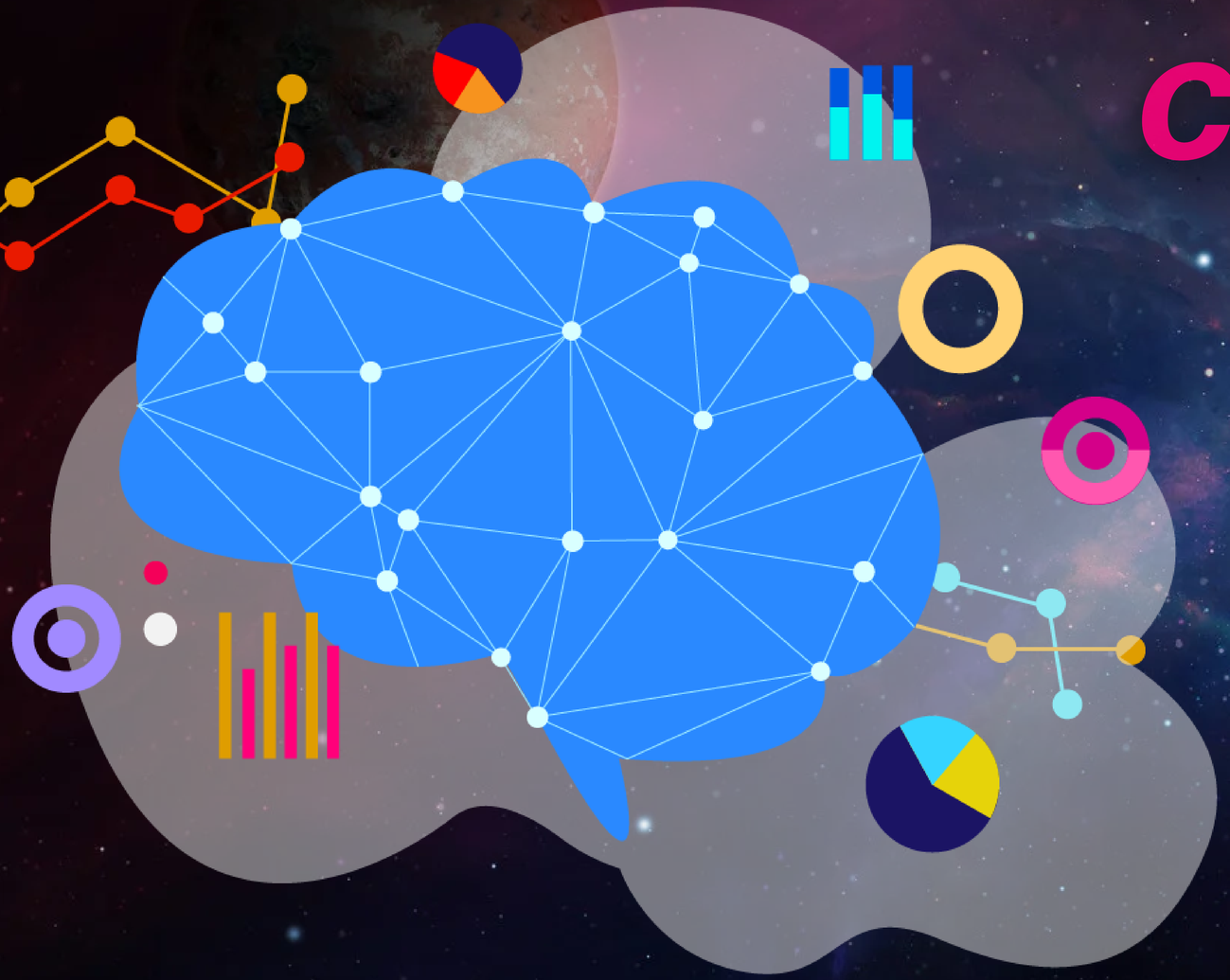
CLASSIFICATION



Classification

Classification Algorithms

- Logistic Regression.
- K-Nearest Neighbours (KNN).
- Support Vector Machine (SVM).
- Decision Tree.
- Naïve Bayes.

etc ...

Logistic Regression

Binary classification example : Fraud or Not Fraud?

We suppose we have a dataset that contains these features :

- Amount of the transaction money.
- Used card.
- Email.
- Signature.

We want to train an algorithm that determines if the transaction is fraudulent or not.

Logistic Regression

Questions

- How do we create a function that returns a value between 0 and 1, where 0=> "Not Fraud" and 1=>"Fraud"?
- How do we calculate how far is our algorithm from giving the right answer?

Answer

- Non-linear functions like Sigmoid function.
- We calculate the cost function.

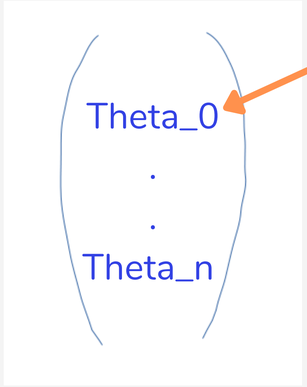
Note

Vectorized Expression of a linear function

f(x0,x1...,xn) = theta_0*x0 + theta_1*x1...+theta_n*xn

is the same as

f(X) = Theta * X

where : Theta = 

Theta_0 = 1

and X = 

b (bias)

Logistic Regression

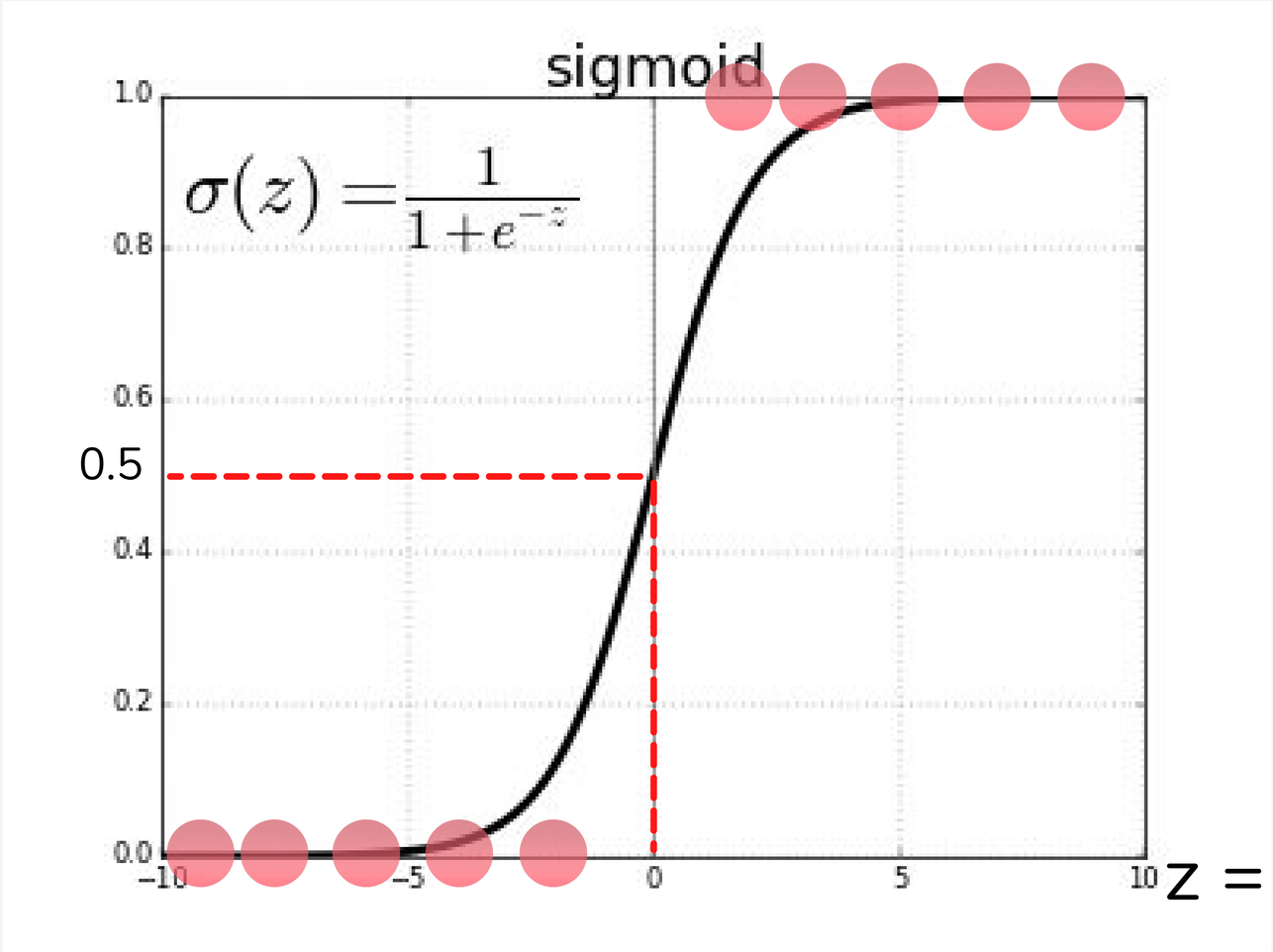
Sigmoid function

z = h(X) = Theta.X

X : vector of features.

Theta : vector of coefficients.

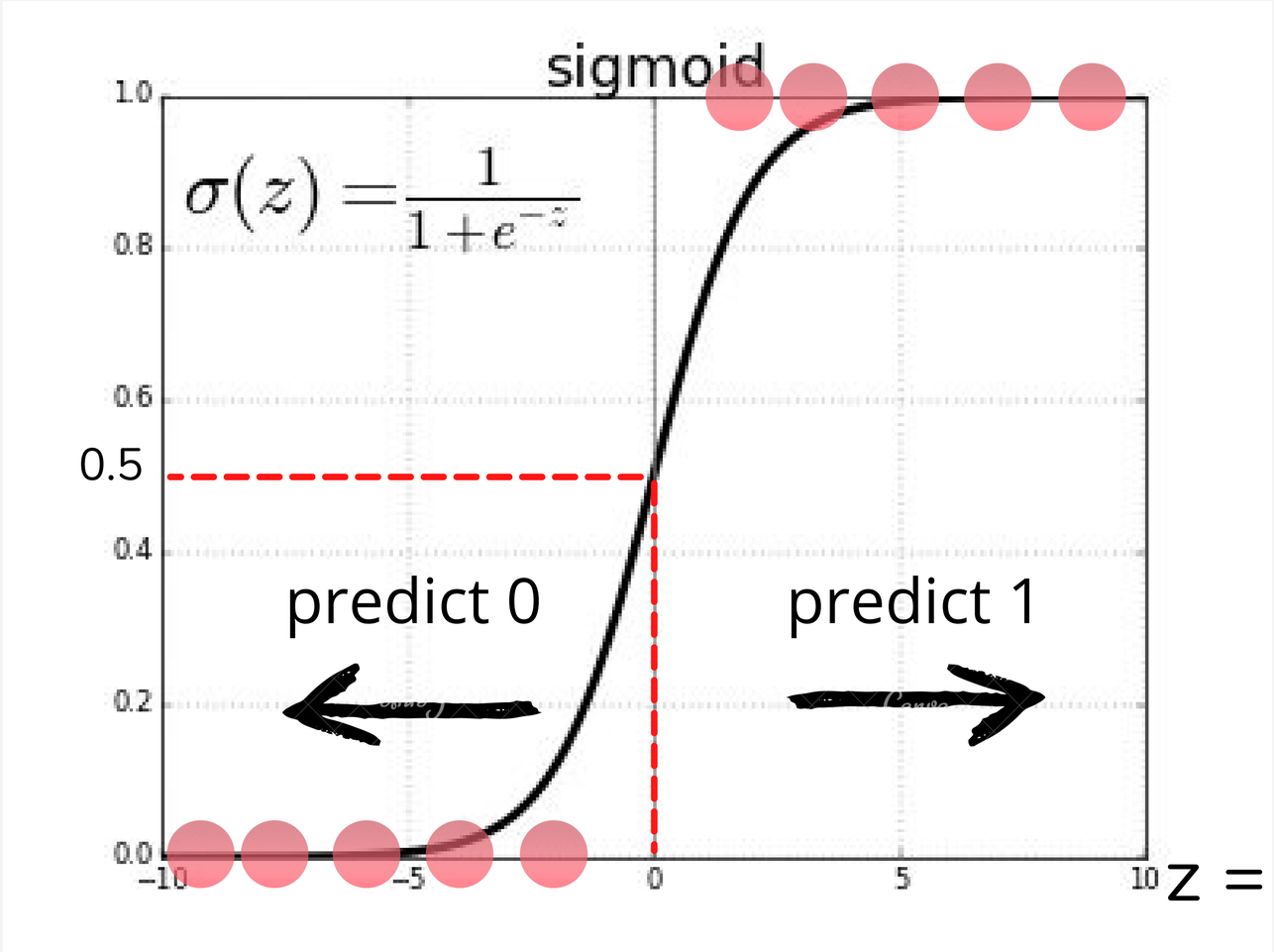
Goal : Find the right Theta that makes sigma(z) fit our data.



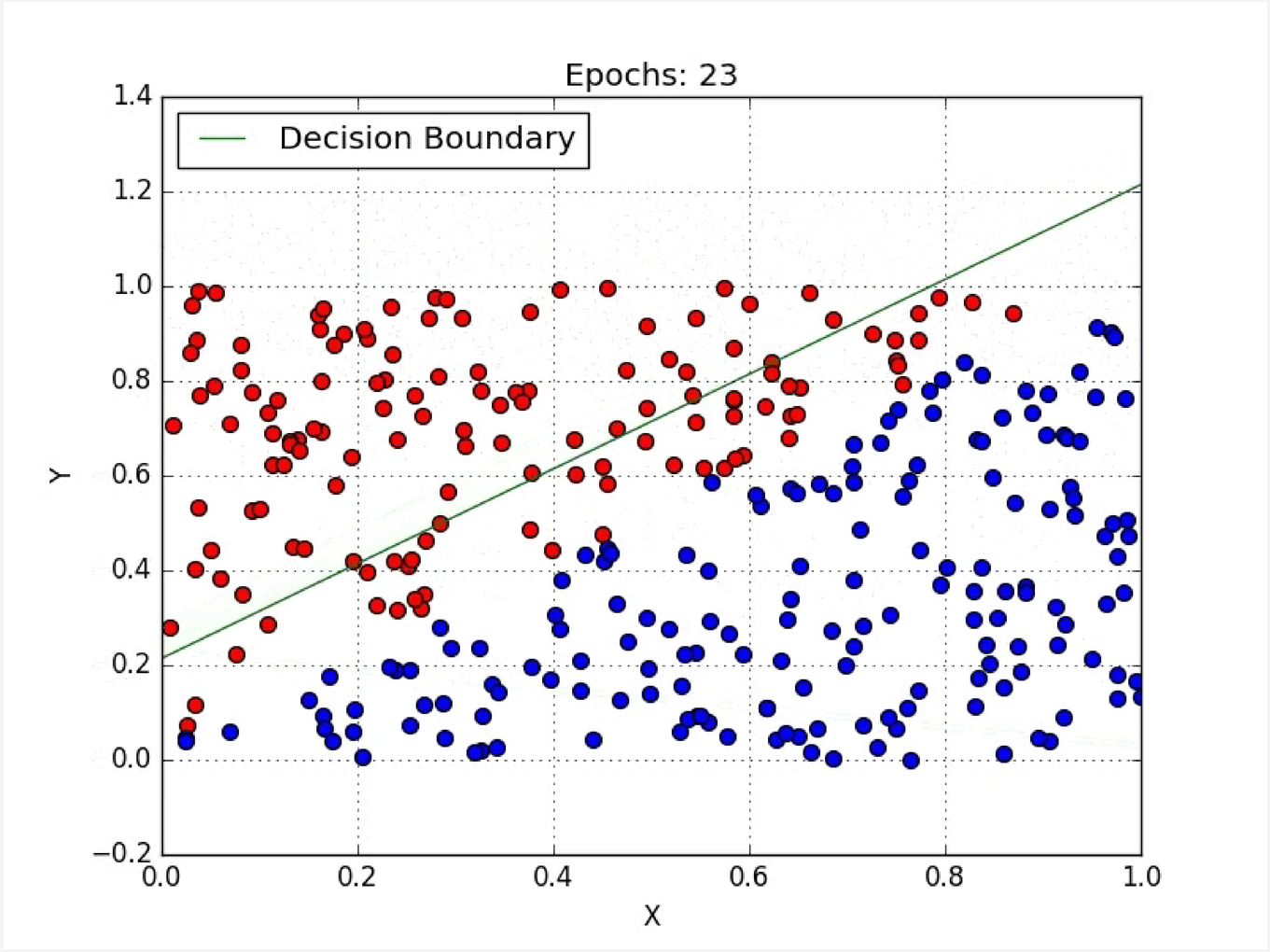
Logistic Regression

if sigma(z) >= 0.5 :
we predict 1
if sigma(z) <0.5 :
we predict 0

Note : we can choose the threshold value (default 0.5).



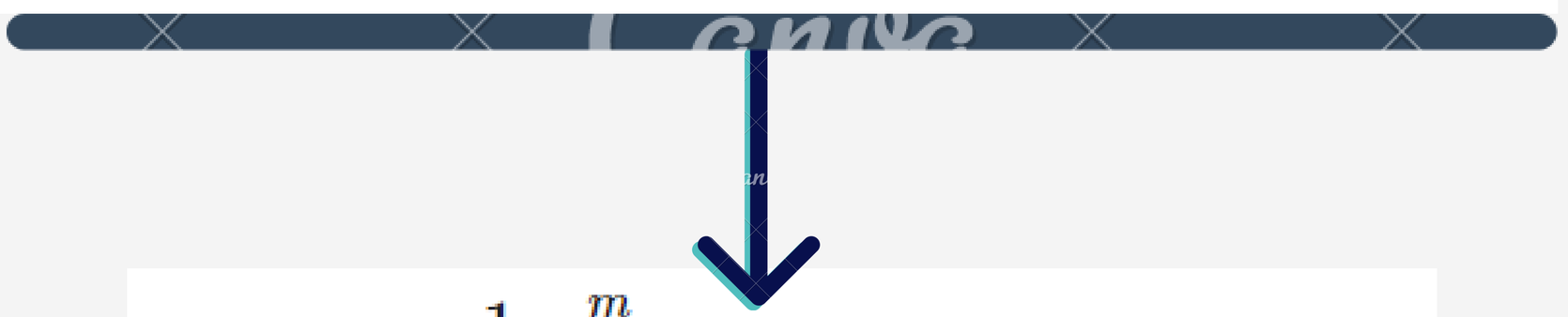
Logistic Regression



Logistic Regression

Cost function for the binary classification problem :

$$J(\theta) = -\frac{1}{m}\sum\left[y^{(i)}\log(h\theta(x(i))) + \left(1-y^{(i)}\right)\log(1-h\theta(x(i)))\right]$$

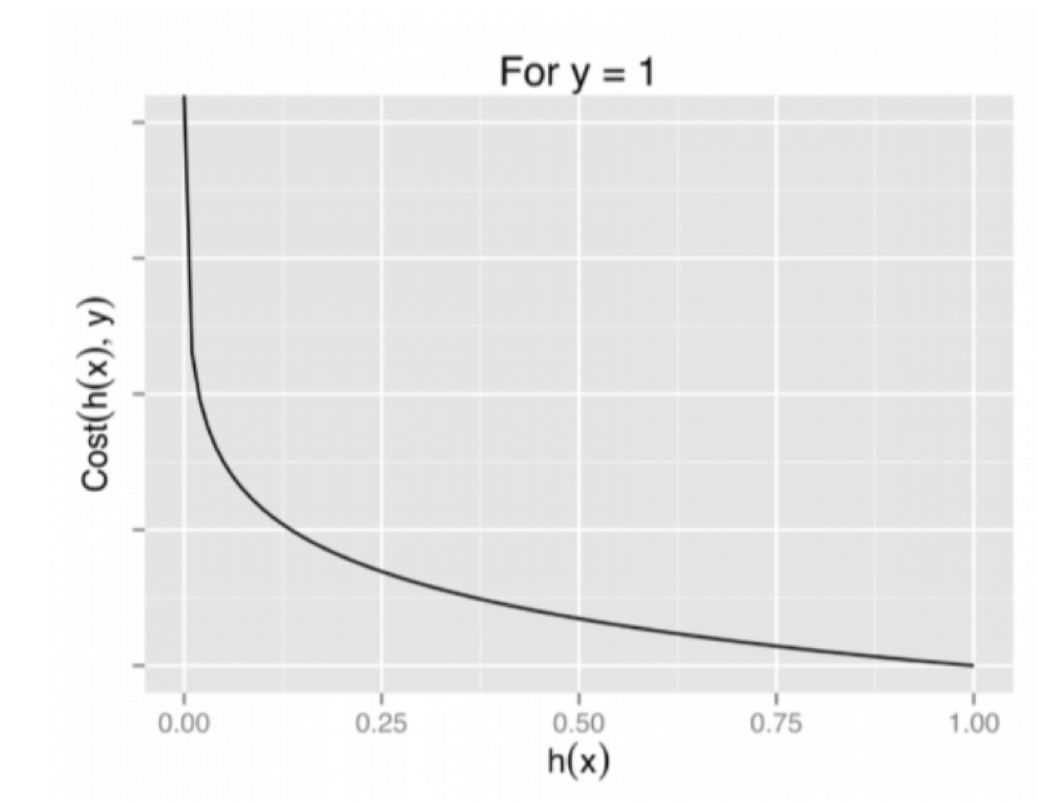


$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\text{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

$$J(\theta) = -\frac{1}{m}\sum\left[y^{(i)}\log(h\theta(x(i))) + \left(1-y^{(i)}\right)\log(1-h\theta(x(i)))\right]$$

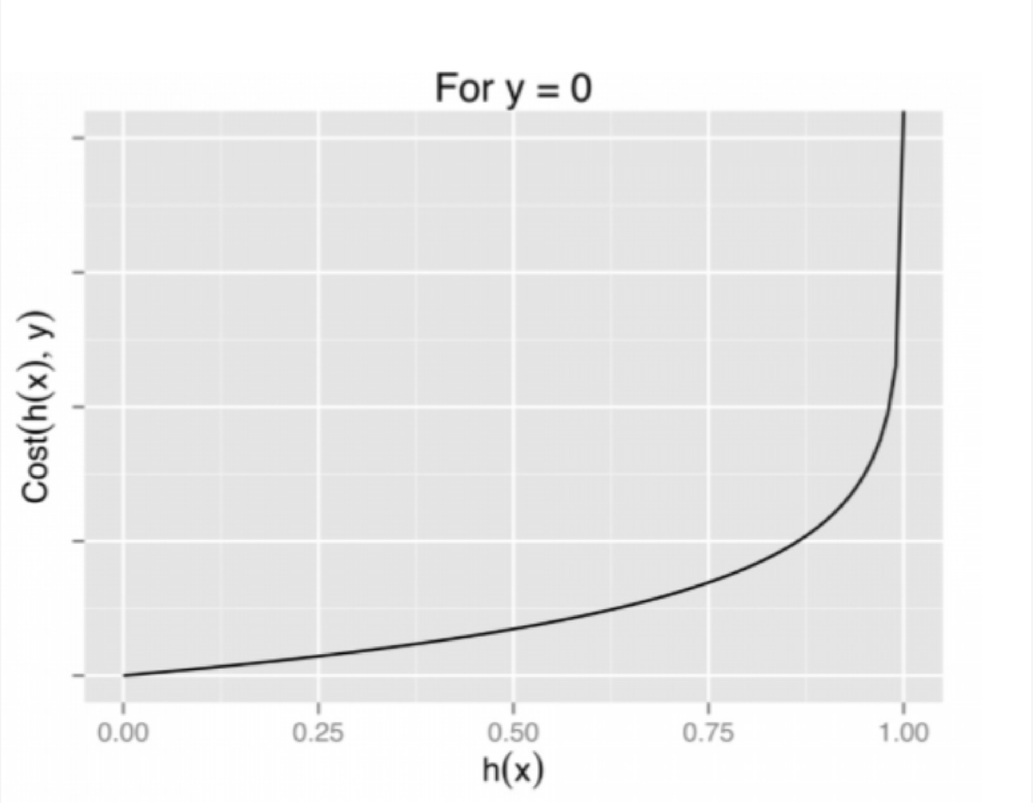
Fraud (yi = 1)

$$\text{Cost}(h_\theta(x), y) = -\log(h_\theta(x))$$



Not Fraud (yi = 0)

$$\text{Cost}(h_\theta(x), y) = -\log(1-h_\theta(x))$$



Let's practice !