Machine Learning

-Ranjith MS

Outline:

- What's Machine Learning??..
- Supervised and Unsupervised Learning
- Hypothesis Function
- Gradient Descent Algorithm
- Feature Scaling and Mean Normalization
- Underfitting and Overfitting
- Back Propagation
- Hands On: Housing Price Prediction

ranjithms523@gmail.com

What’s Machine Learning or Artificial Intelligence and Where is it used???.....

- web search
- Facebook or Apple’s recognizes your friends in your picture
- Spam filter in emails
- self-driving cars
- Gaming Control

Neural Style Transfer

Content Image

Style Image

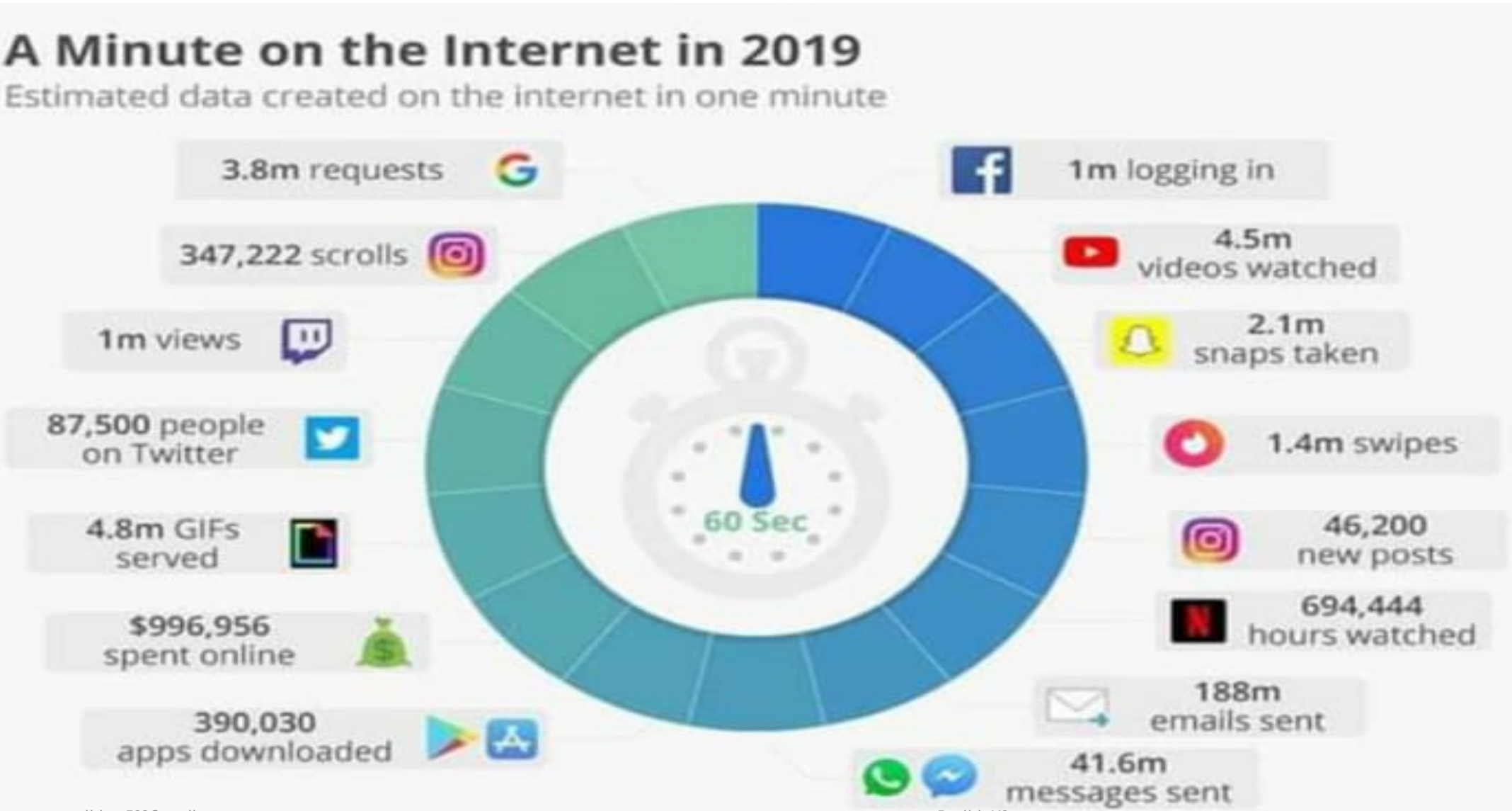
Generated Image

Content Image

Style Image

Generated Image

Face recognition system



ranjithms523@gmail.com Ranjith MS


HTML + CSS + JS + PHP
= Mode 1
HTML5 + Bootstrap + React JS + Laravel
= Mode 2
Mode 2 + Machine
Learning + Deep Learning

Question: What is the Relation between X and Y ??...

X	Y
2	2
-3	-3
5	5
6	6
0	0
1.3	1.3
10	10

- X=Y

ranjithms523@gmail.com Ranjith MS

Question: What is the Relation between X and Y ??…

X	Y
2	-4
3	-6
-2	4
5	-10
0	0

- Y=-2X

ranjithms523@gmail.com

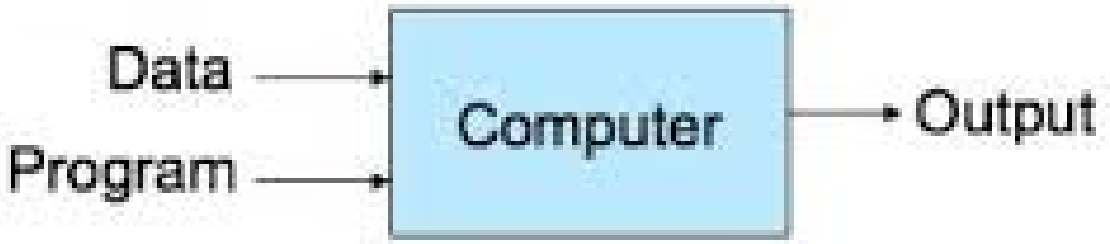
Now!!…
What's the relation between X and Y???…

X	Y
2	4
7	19
-5	-17
10	28
15	43

- Y=3X-2

ranjithms523@gmail.com

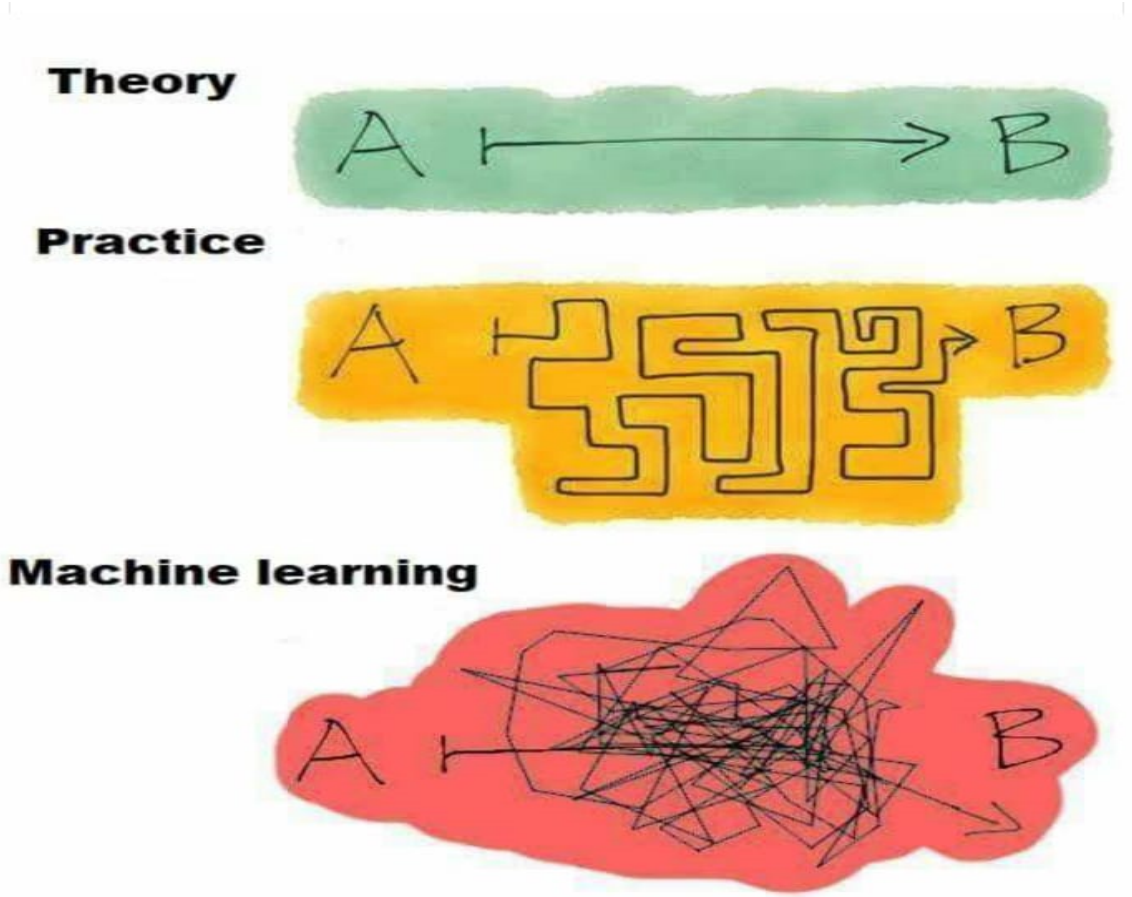
Programming



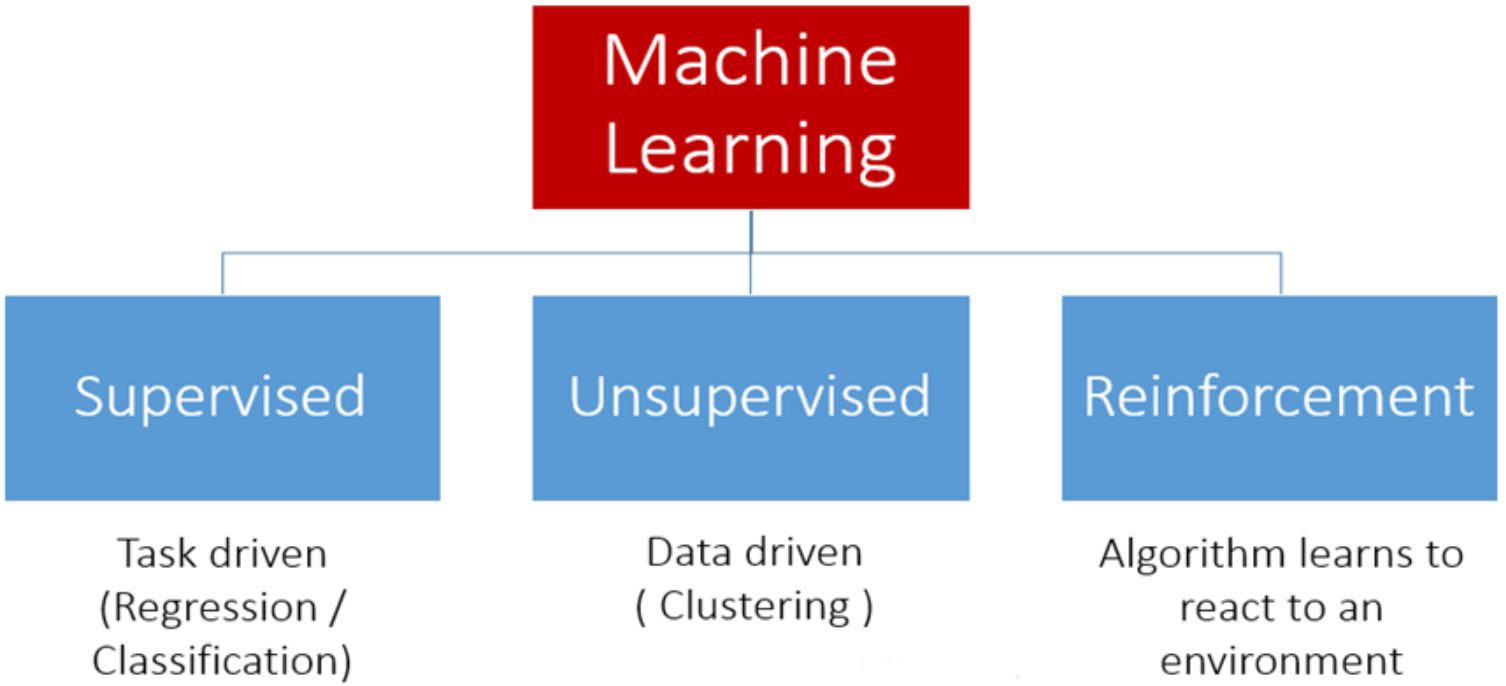
Machine Learning

So What's Machine Learning??...

- **Machine learning is a sc getting computers to le without being explicitly programmed.**



Types of Machine Learning



Supervised learning

Supervised learning is where you have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output.

Eg : Predicting the house price given the size of the house based on the previous data.

Classification of Supervised Learning

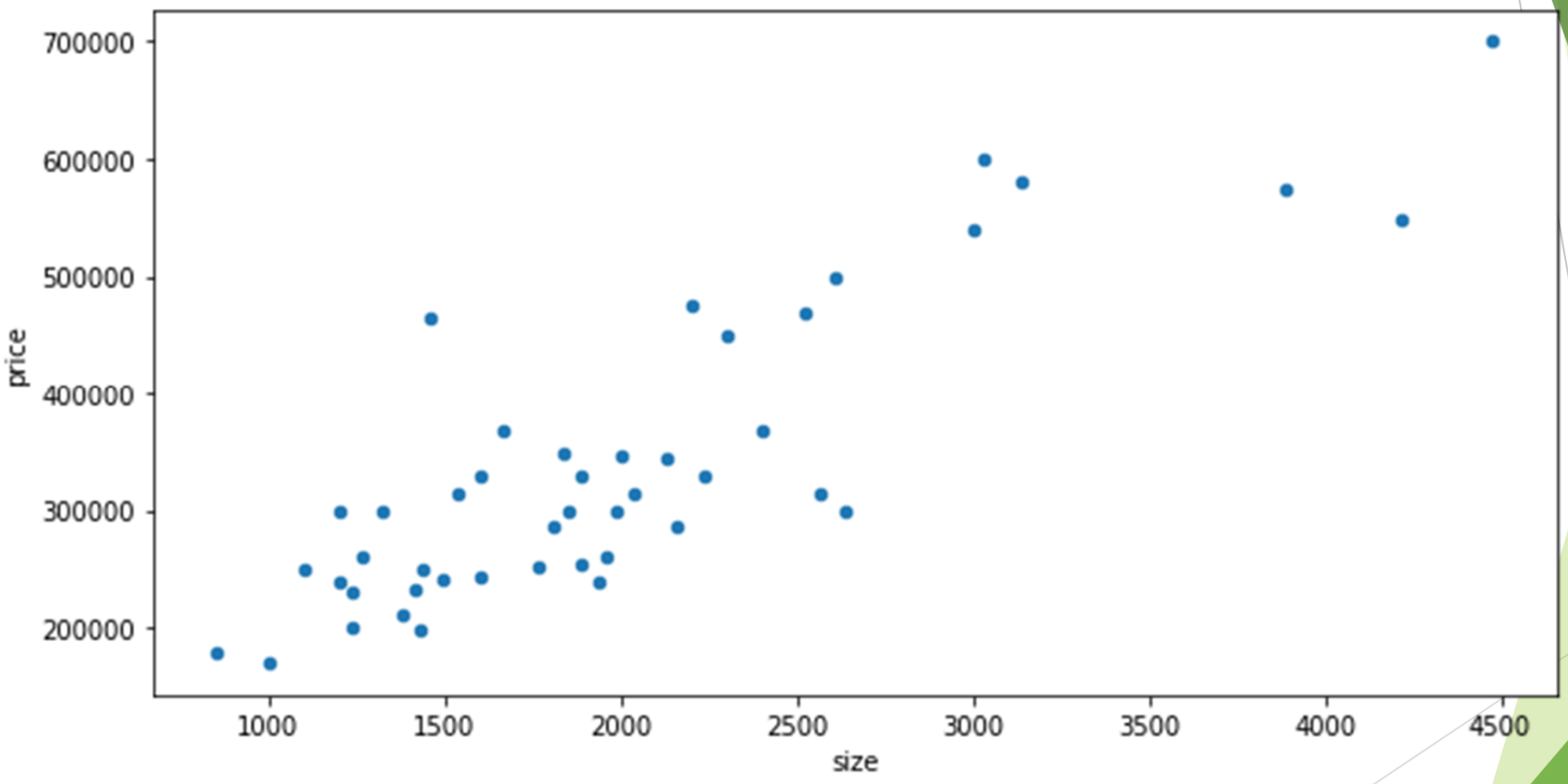
Regression

- A regression problem is when the output variable is a real value, such as “dollars” or “weight”.
- Eg: Given a picture of a person, we have to predict their age on the basis of the given picture

Classification

- A classification problem is when the output variable is a category, such as “red” or “blue” or “disease” and “no disease”.
- Eg: Given a patient with a tumor, we have to predict whether the tumor is malignant or benign.

Regression


700000
600000
500000
400000
300000
200000
price
1000
1500
2000
2500
3000
3500
4000
4500
size

Training Set:

Notation:

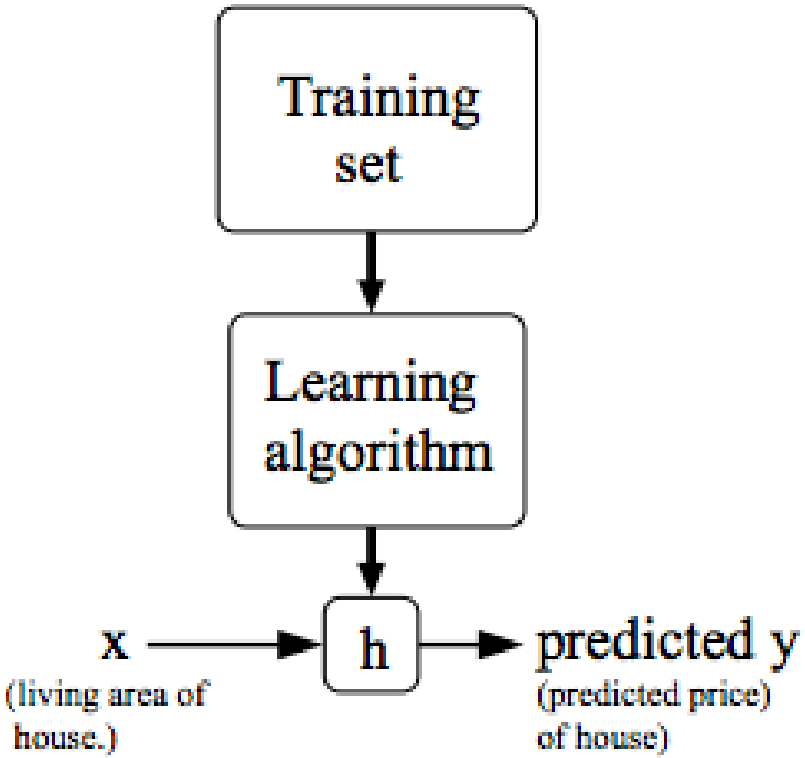
m=Number of Training Examples

x's = "input" variable / features

Y's = Output/Target Variables

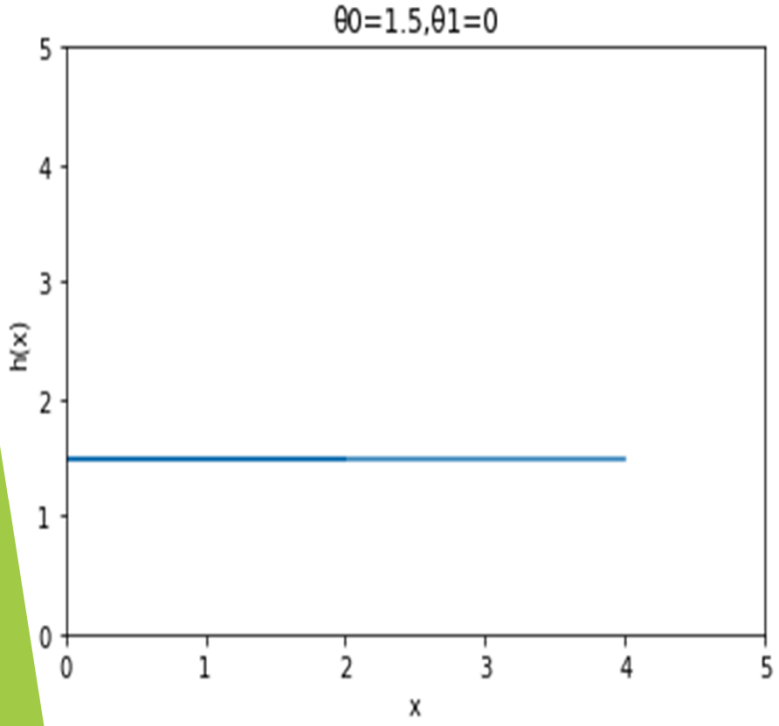
Hypothesis Function

$$h_\theta(x) = \theta_0 + \theta_1 x$$

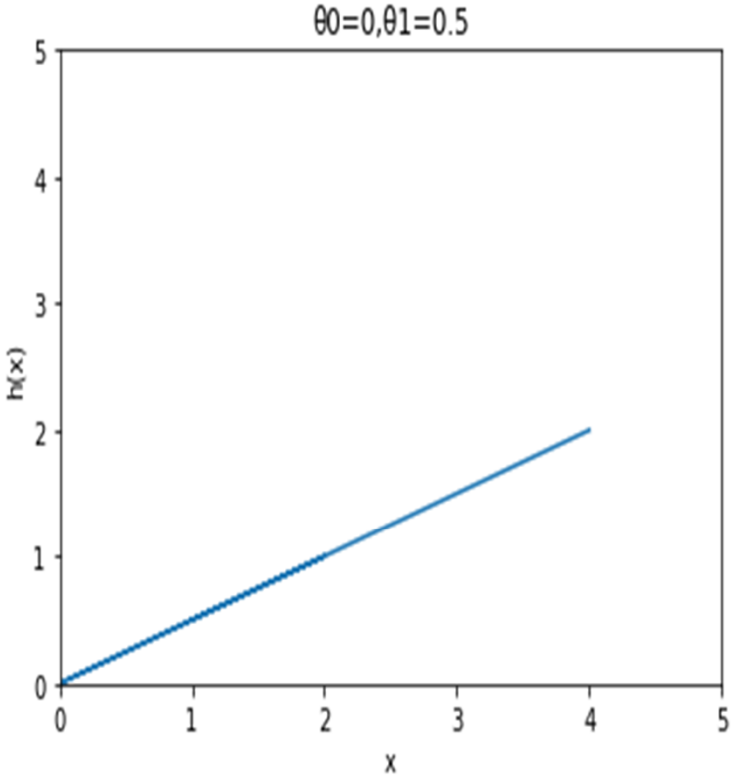


How To Choose Parameters???....

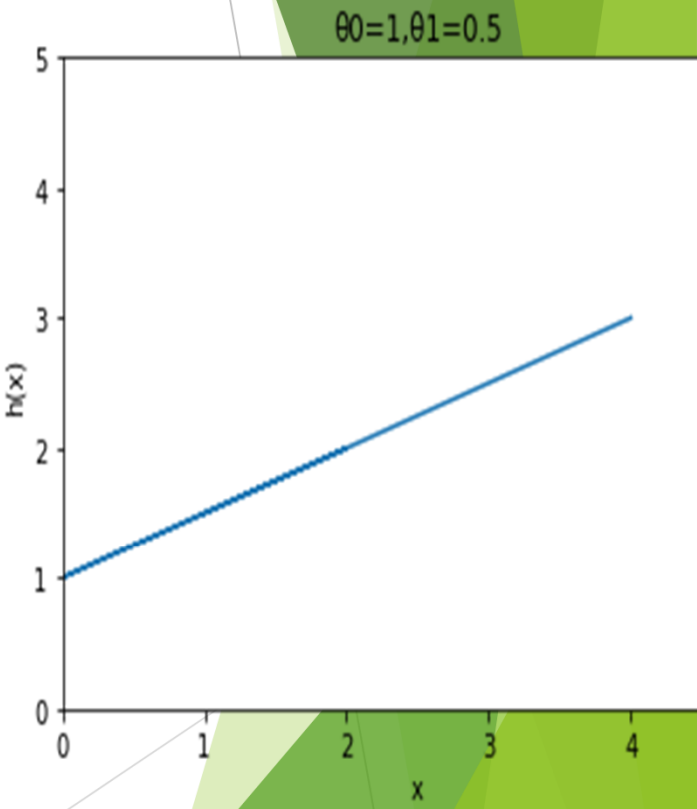
$\Theta_0=1.5, \Theta_1=0$



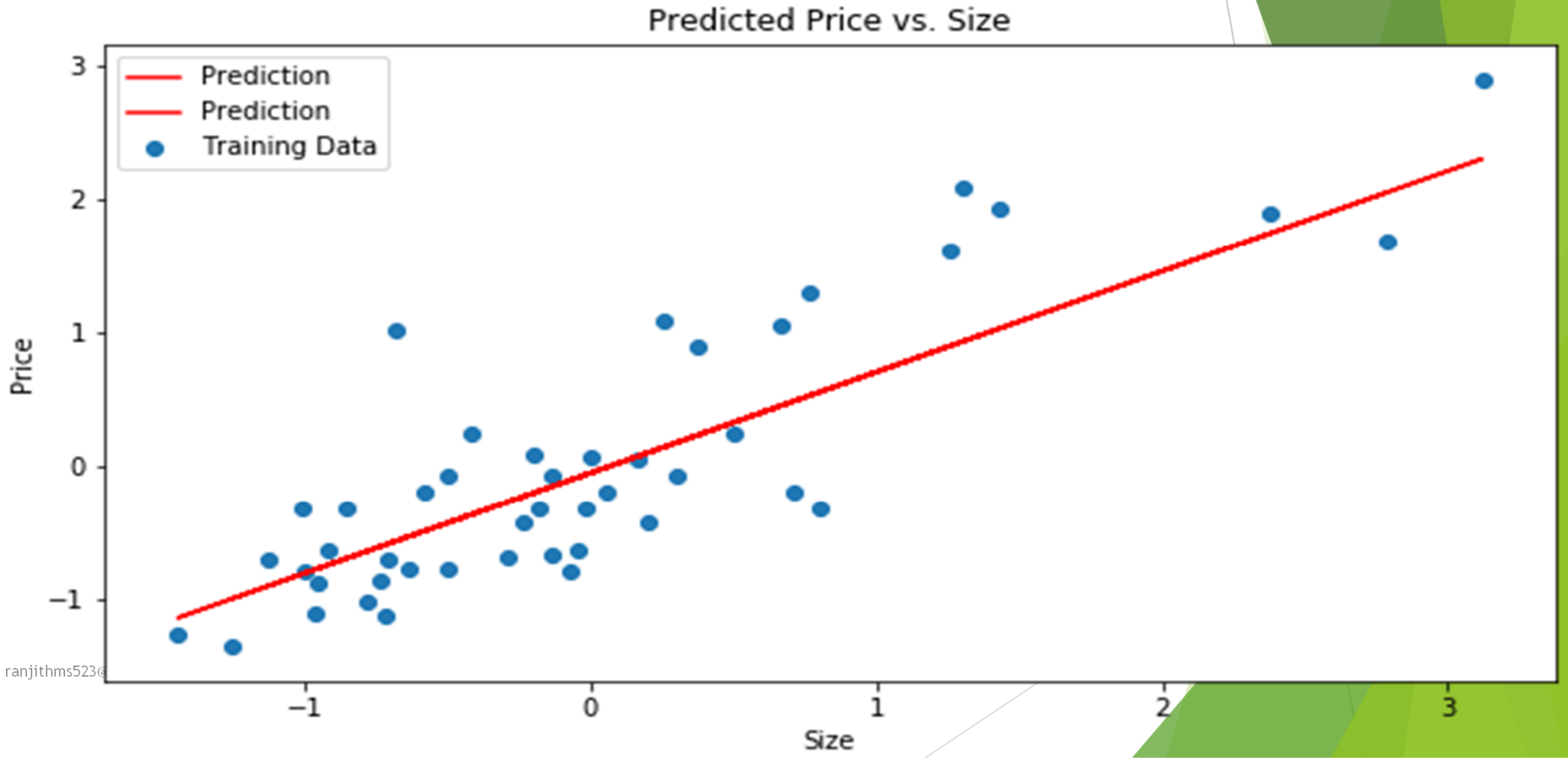
$\Theta_0=0, \Theta_1=.5$



$\Theta_0=1, \Theta_1=.5$

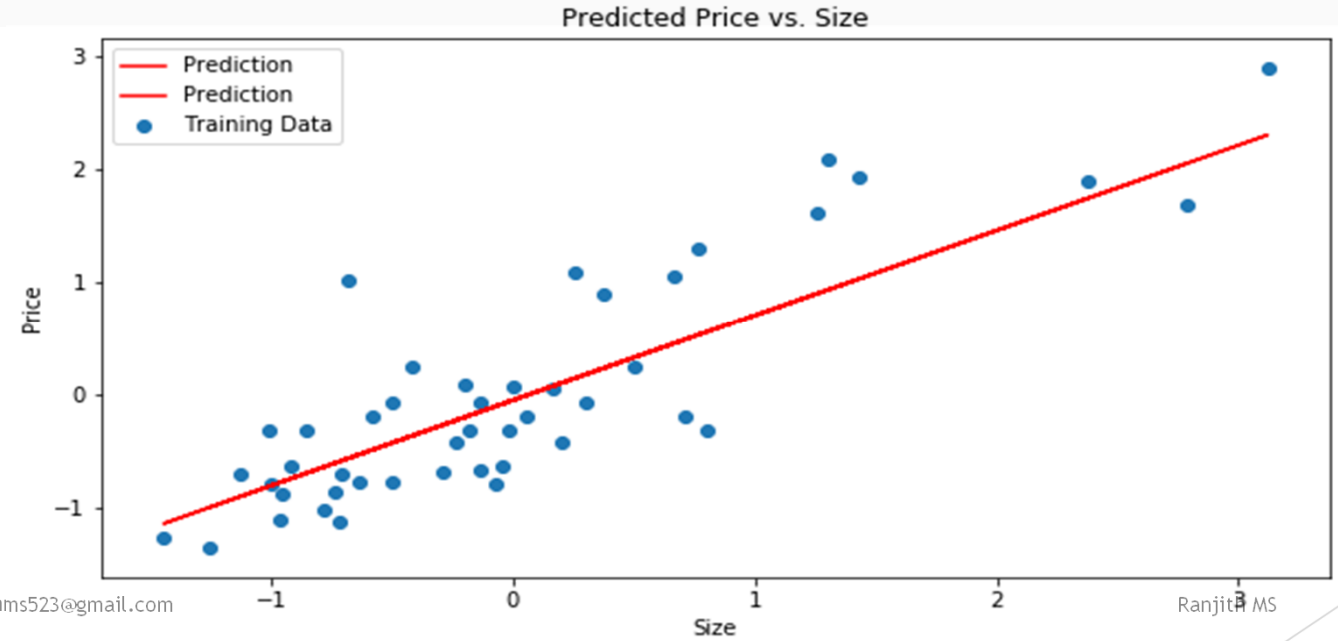


Choose parameters so that h(x) is close to y for training data



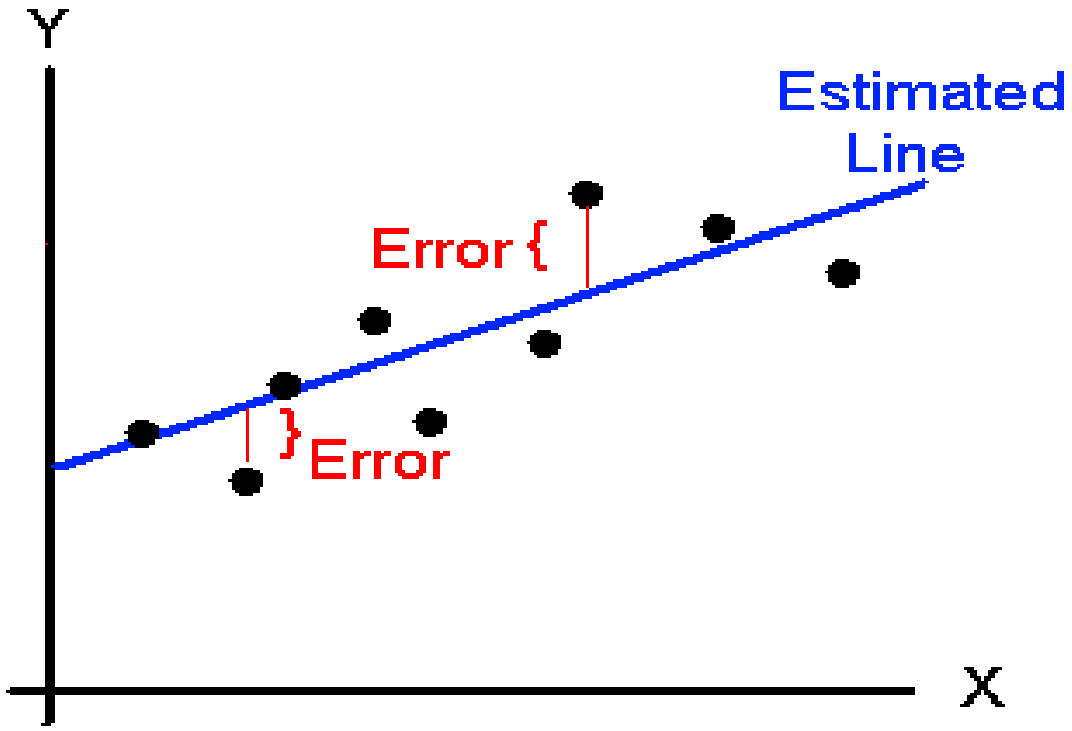
Cost function [Squared error function]

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$



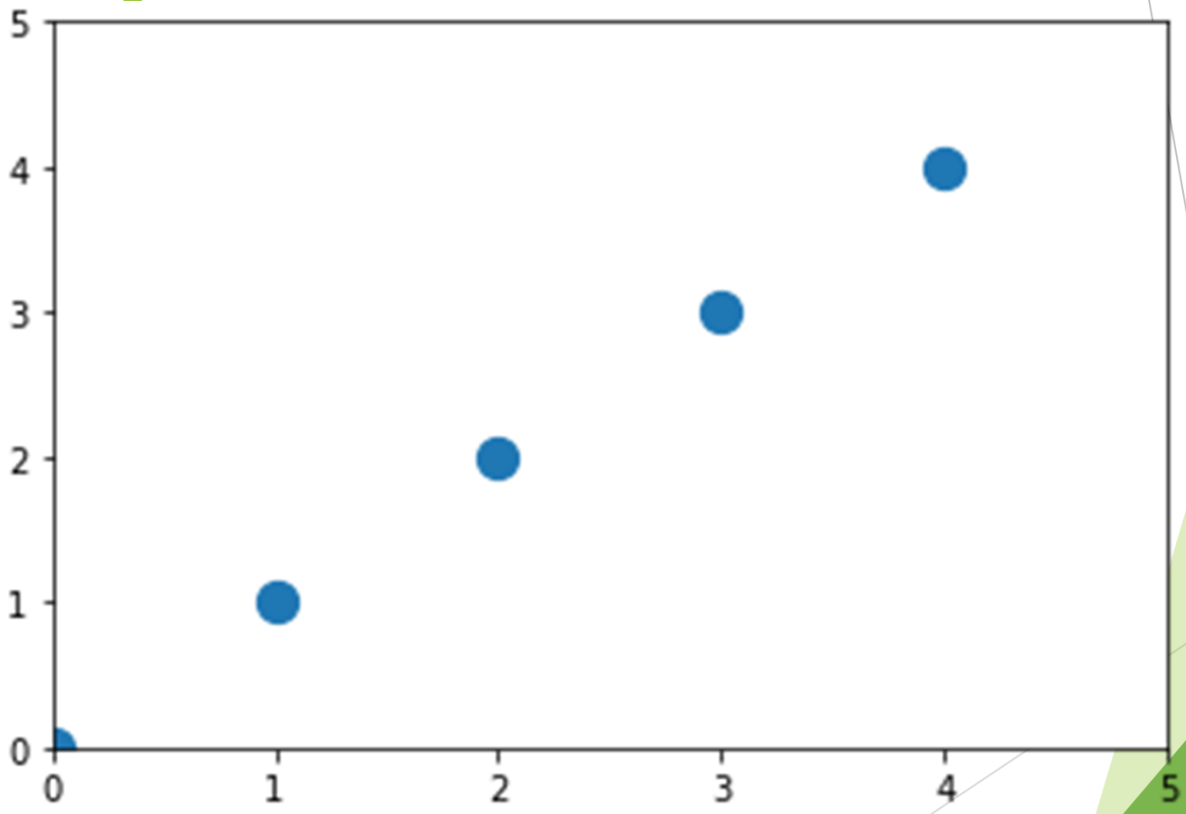
ranjithms523@gmail.com

Ranjith MS

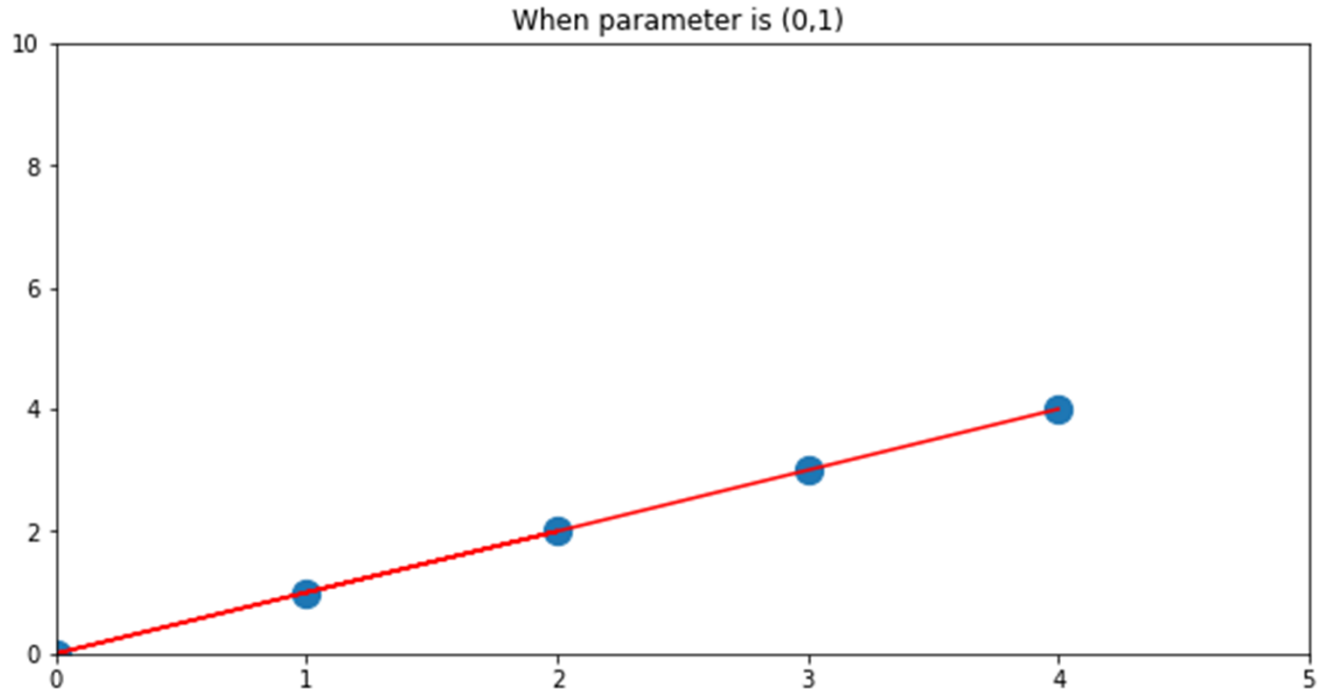


Cost / Loss Function

Let x=[0,1,2,3,4]
y=[0,1,2,3,4]

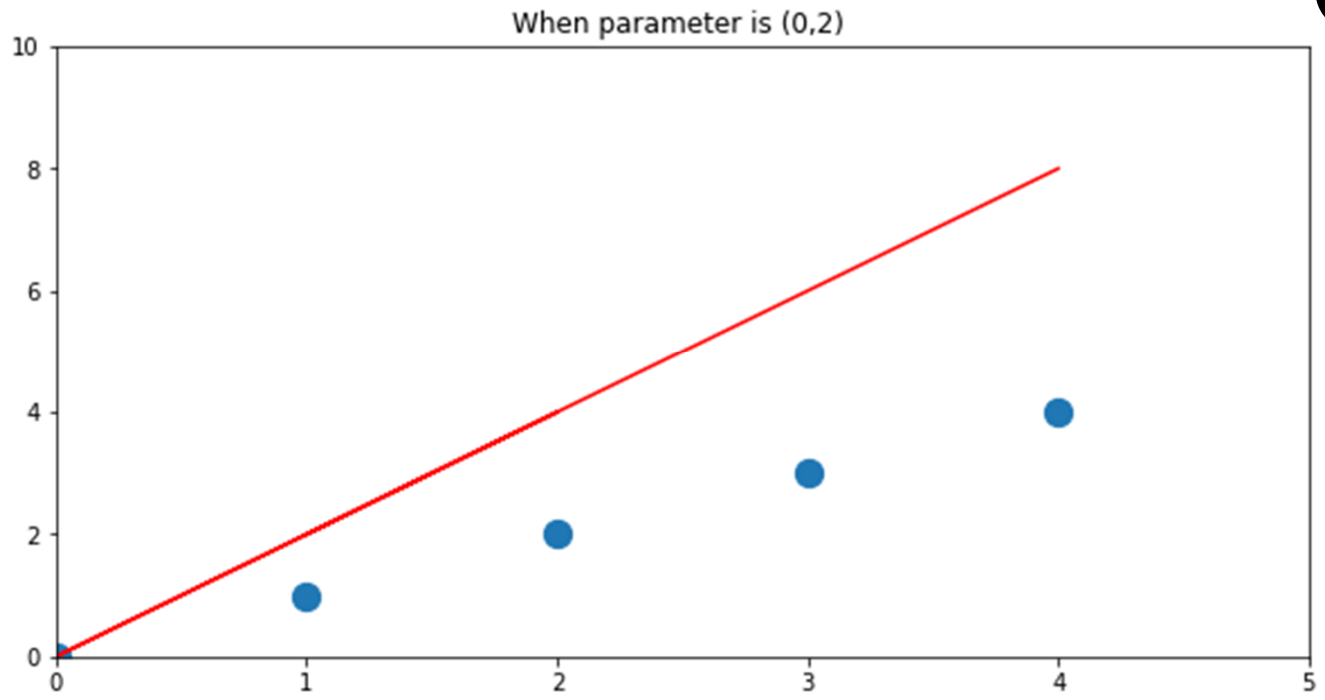


Let us assume $\Theta_0=0$
What is the cost when $\Theta_1=1$???....



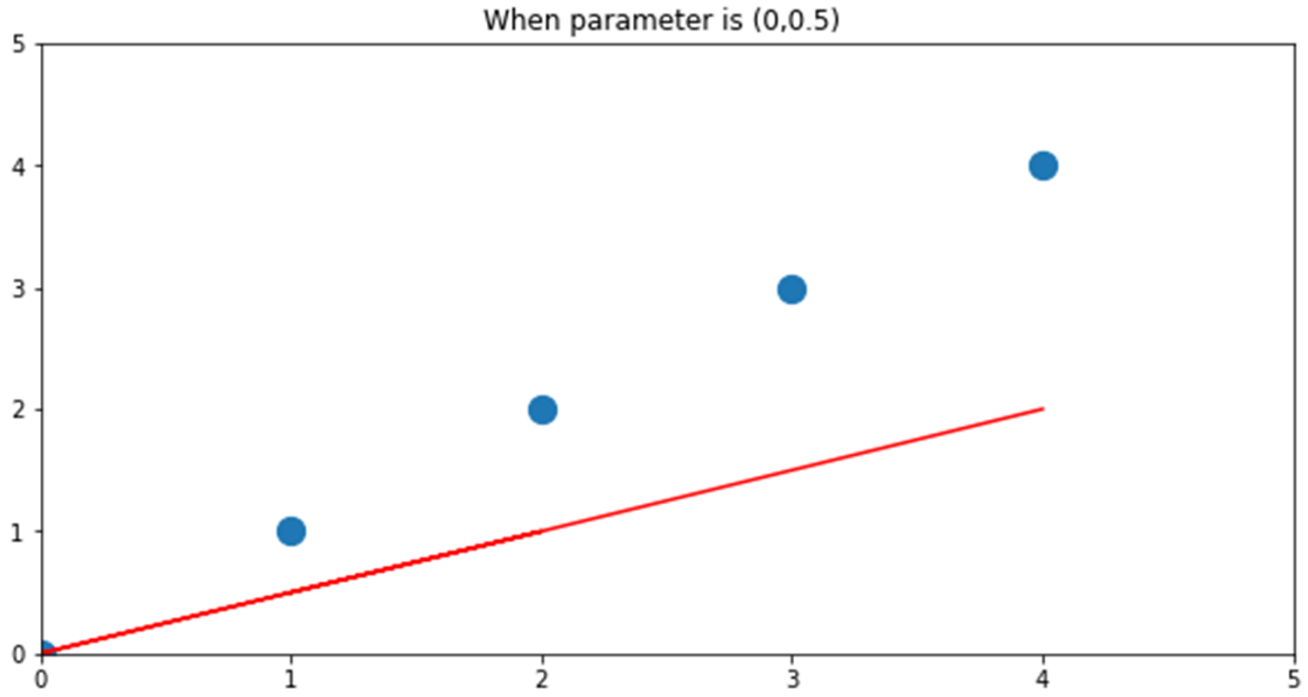
Cost = 0

What is the cost when Θ_1=1 ???....



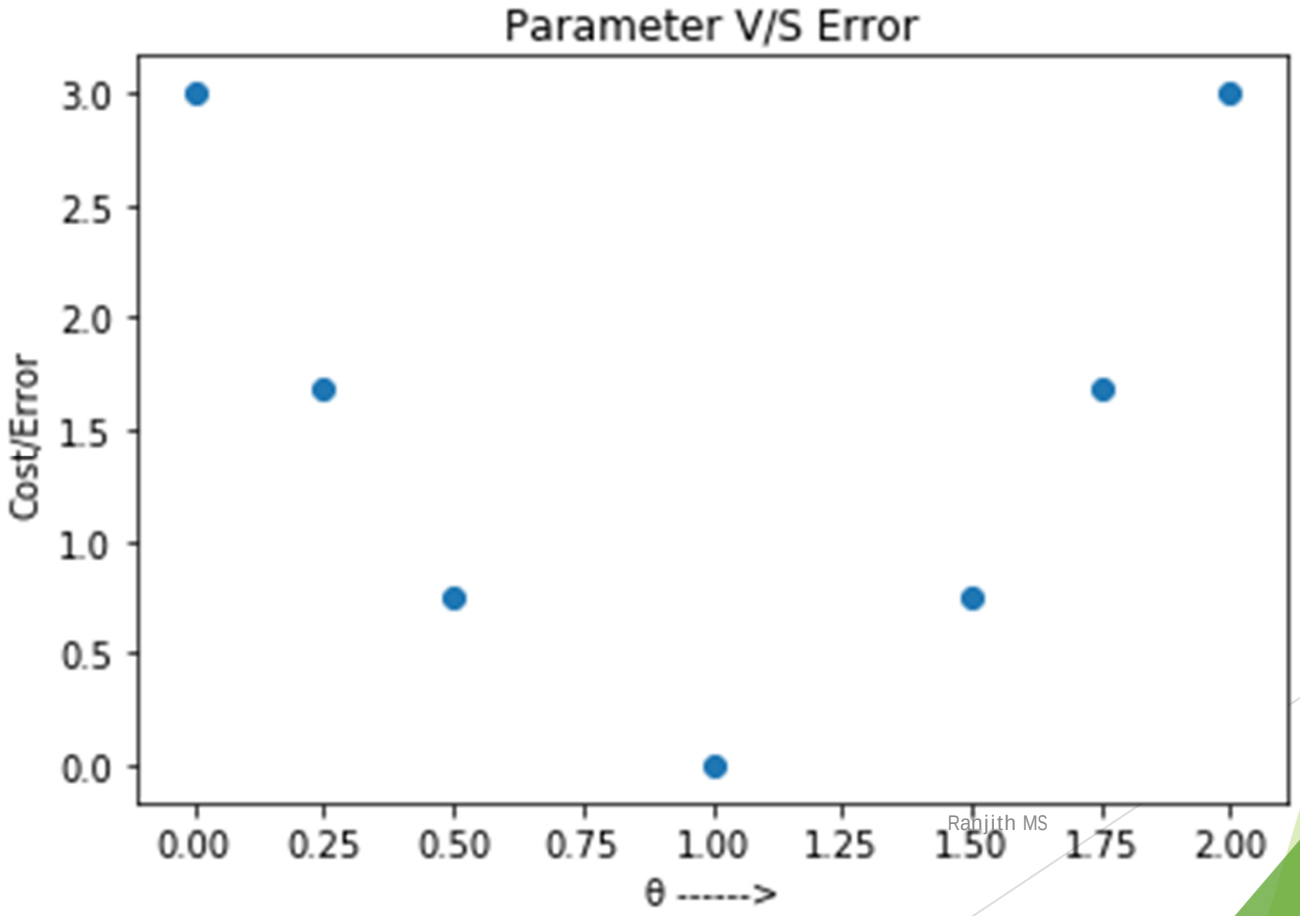
Cost = 3.0

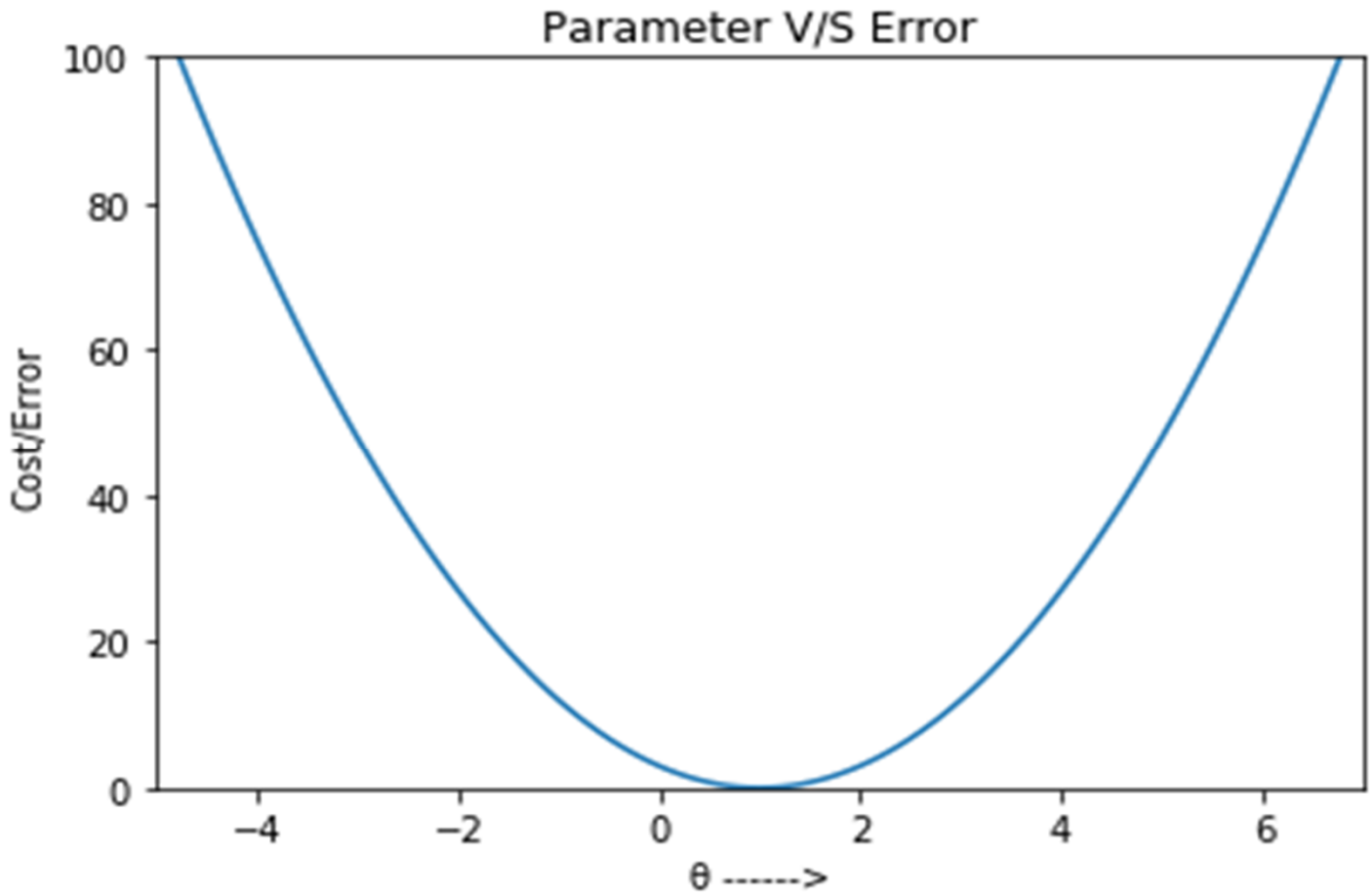
What is the cost when Θ_1=0.5 ???....



Cost =0.75

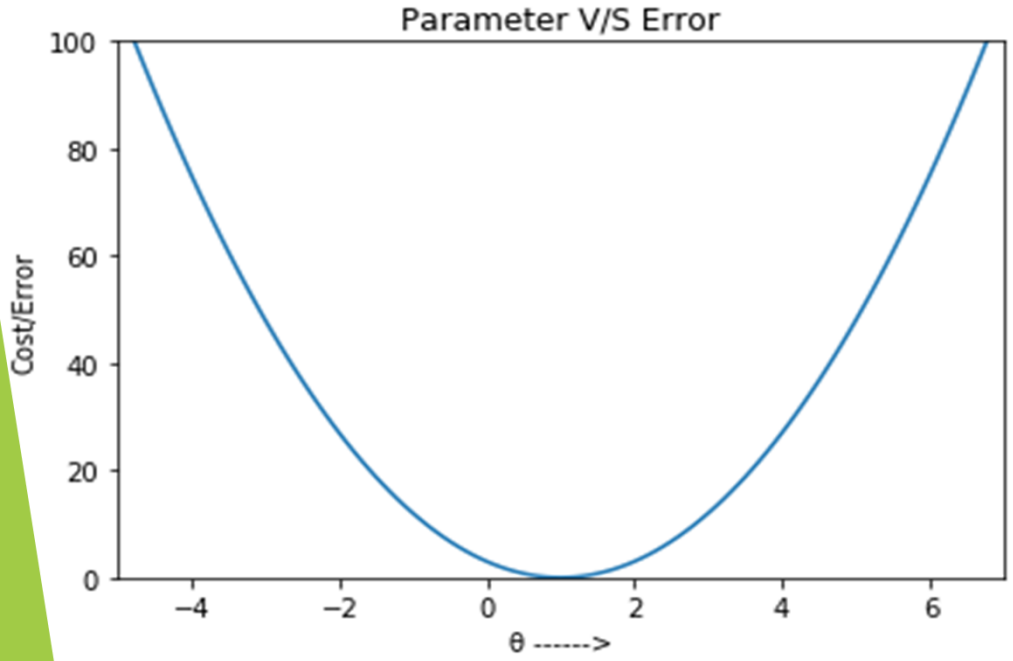
How Error varies with θ

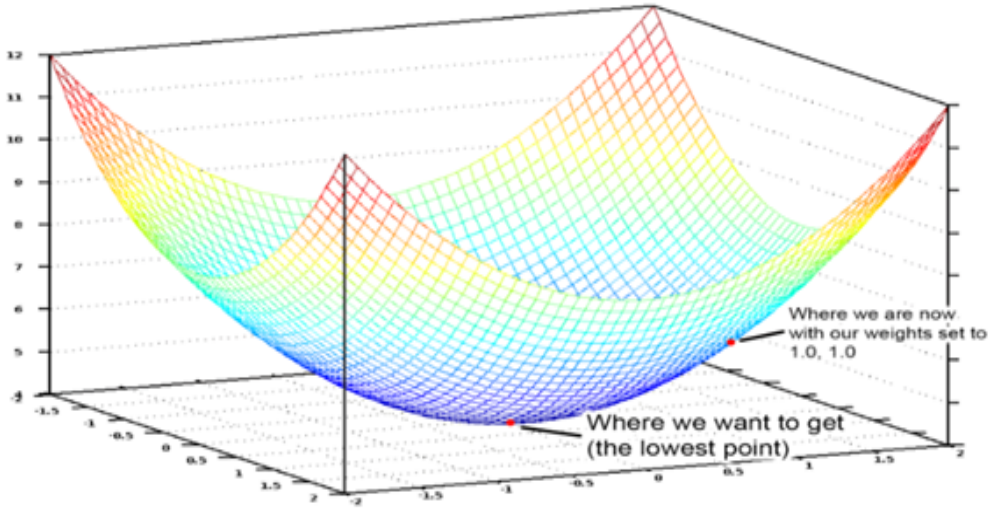



Parameter V/S Error
100
80
60
40
20
0
Cost/Error
−4
−2
0
2
4
6
θ ------>

Gradient Descent

- It is a method to find the values of Θ such that J(Θ) will be minimum





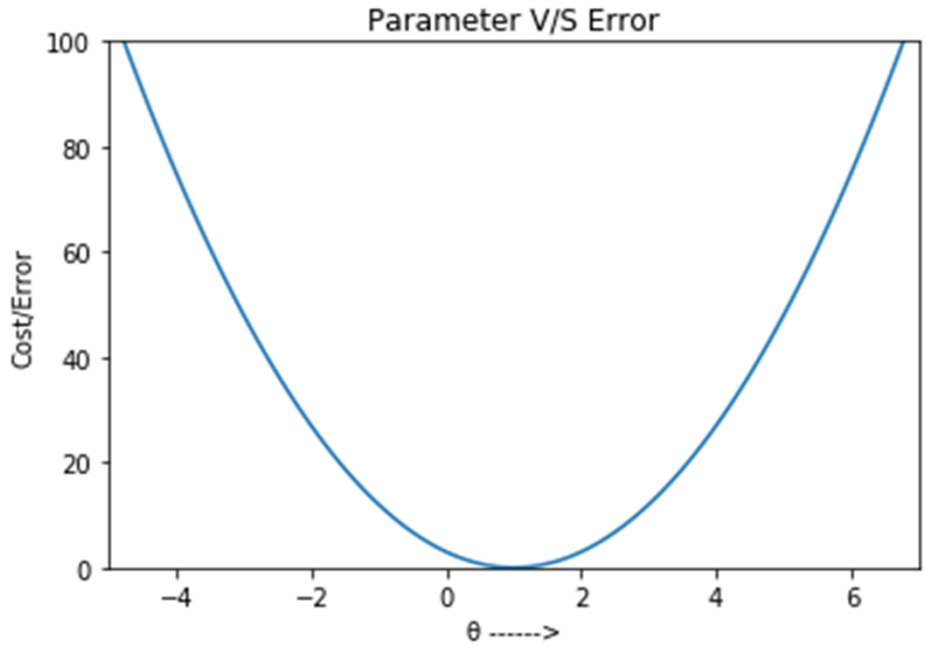
Gradient Descent

- What is the derivate of an increasing Function ??....
- Positive

- What is the derivate of an decreasing Function ??....
- Negative

Gradient descent

- Initialize the weights randomly
- Update the weights based on the derivative (i.e, derivative of cost function w.r.t parameter)
- Decrease the initialized/previous weight if derivative is positive
- Increase the initialized/previous weight if the derivative is negative
- Continue the process until you reach minima

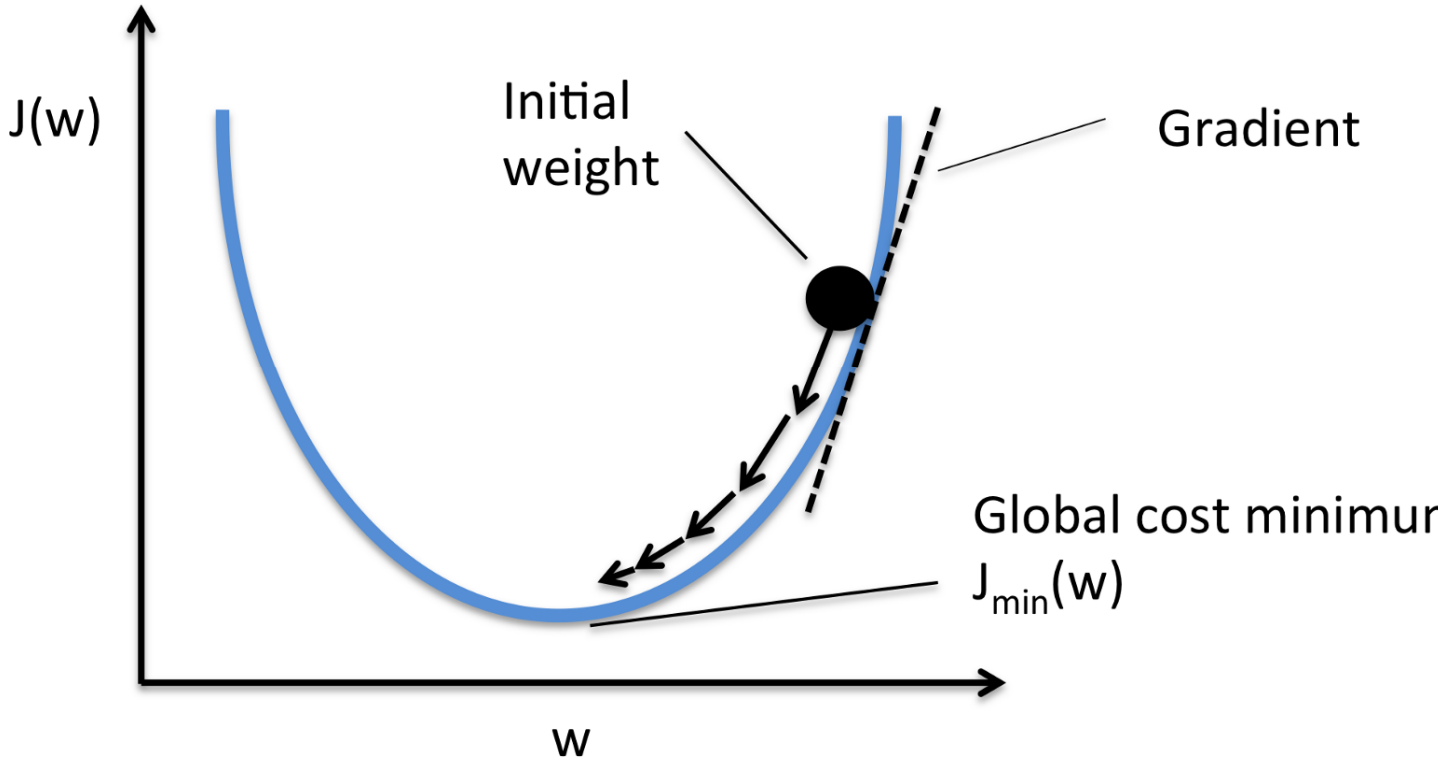


Gradient descent

Repeat until convergence {

$$\theta_j \leftarrow \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

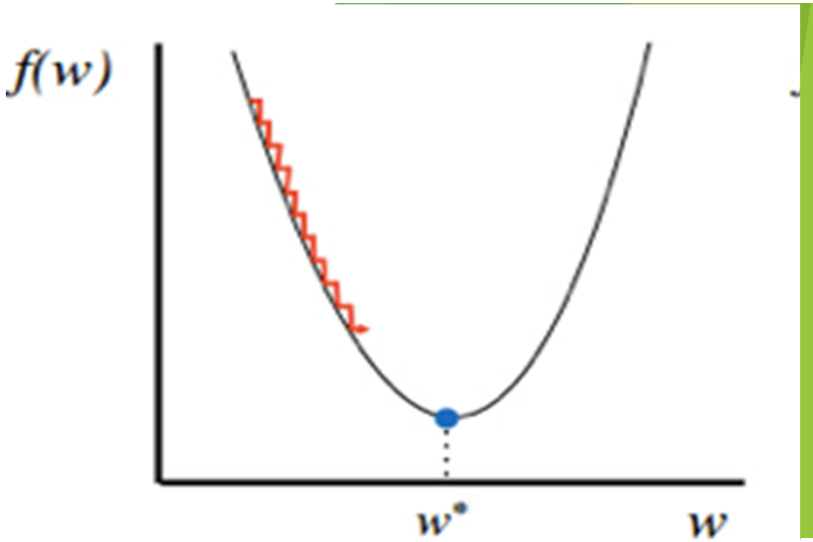
}

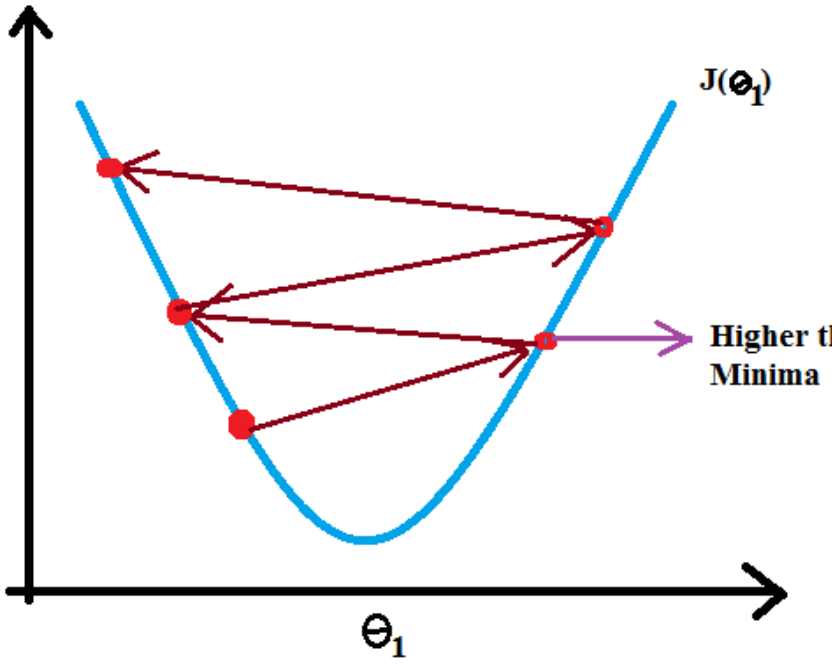


The size of each step is determined by the parameter α, which is called the learning rate.

Debugging gradient descent

- If α is too small then the gradient descent can be slow
- If α is too large then the gradient descent may overshoot the minima and may fail to converge

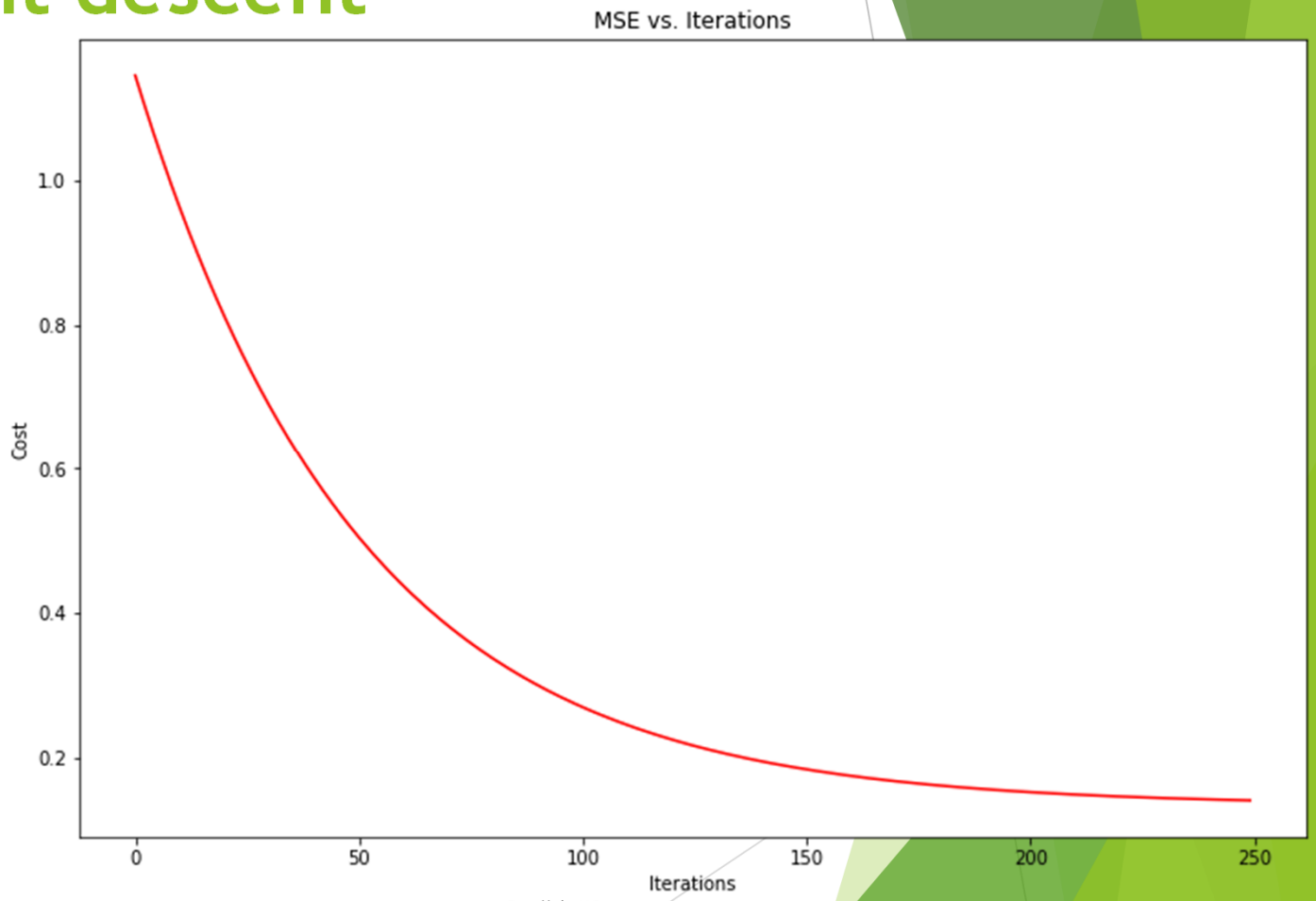




ranjithms523@gmail.com

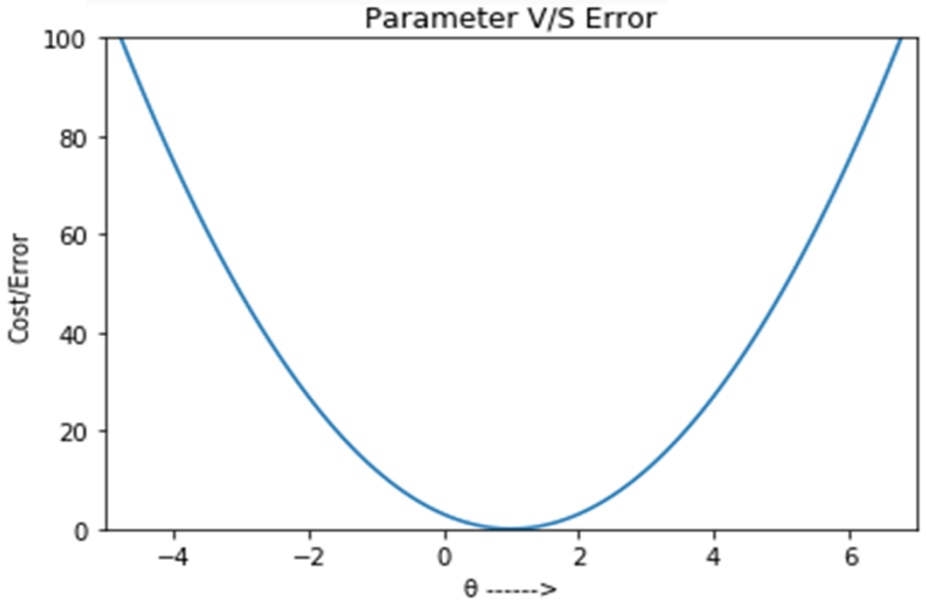
Debugging gradient descent

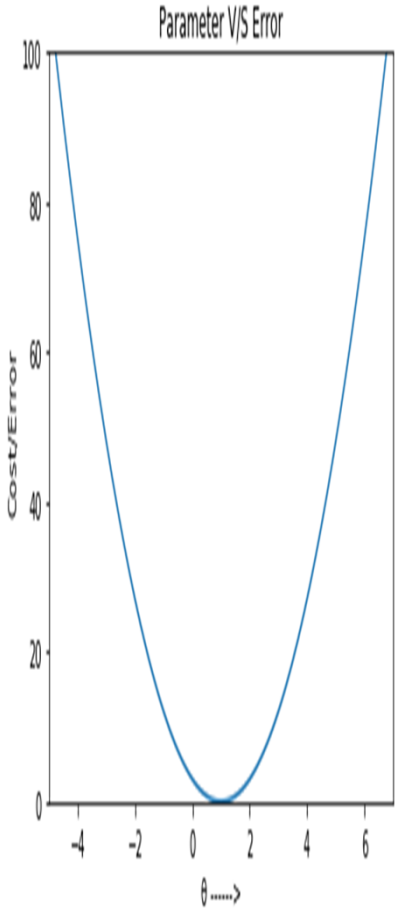
- Now plot the cost function, J(θ) over the number of iterations of gradient descent. If J(θ) ever increases, then you probably need to decrease α.



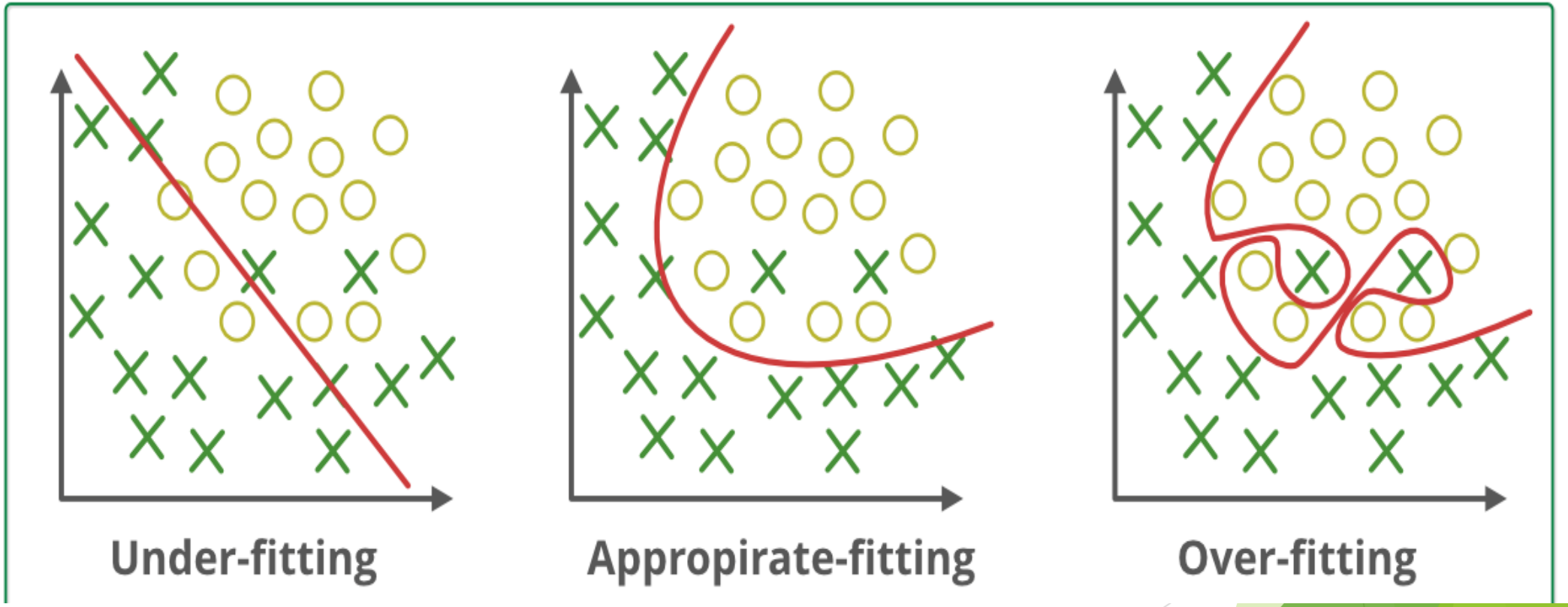
Feature Scaling

$$x_i := \frac{x_i - \mu_i}{s_i}$$



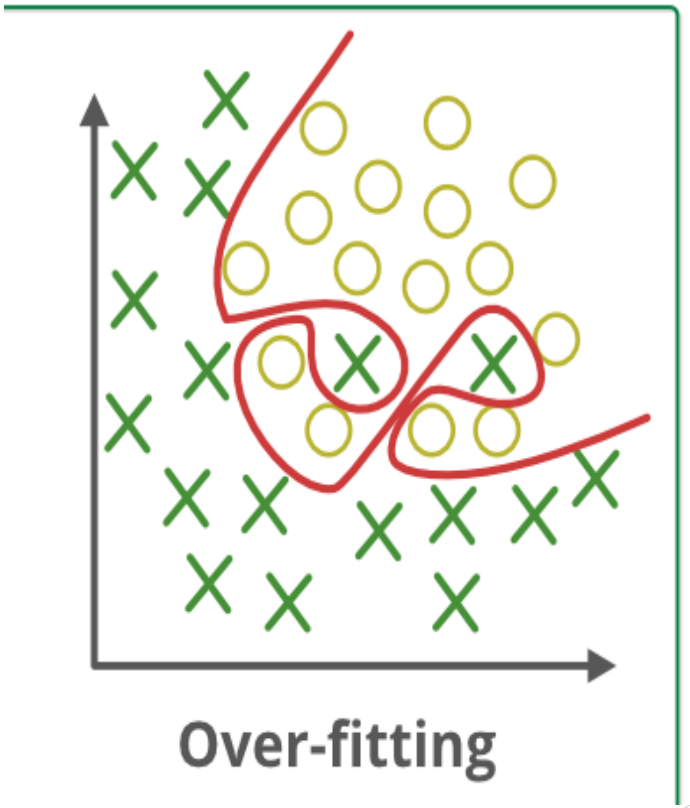


How to choose the curve??.....

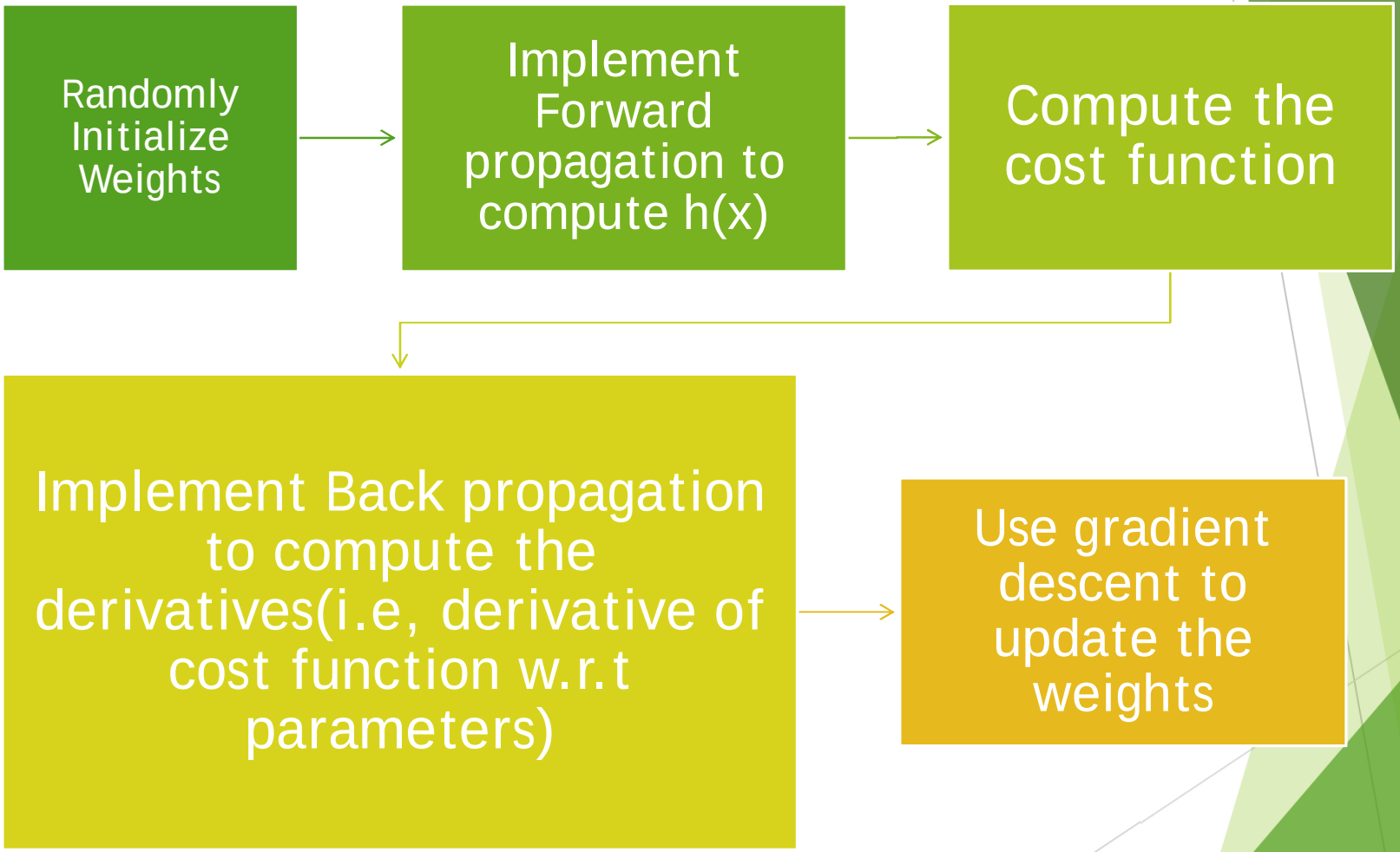


Ways to reduce overfitting

- Dropout
- Regularization

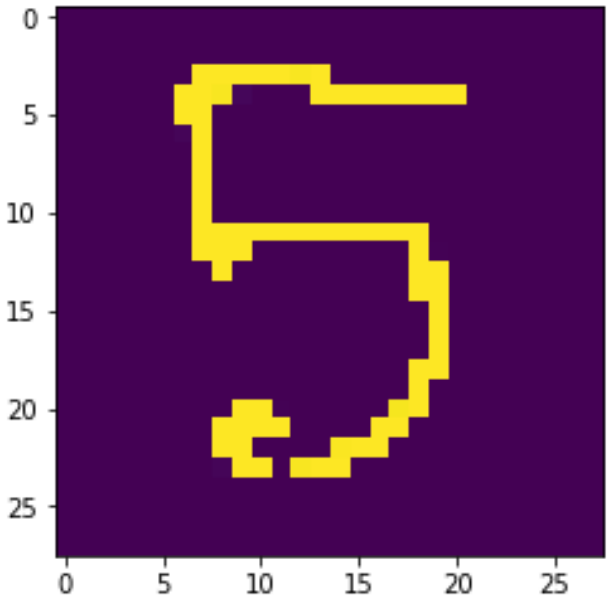


Training a model



Coming up.....
Digit Recognition

Predicted value is 5



Predicting from the model

```
[ ]  #path="ML workshop/con_3.jpg"
     img = image.load_img(path,color_mode = "grayscale",targ
     img = image.img_to_array(img)
     test_img = img.reshape((1,784))
     ans=model.predict(test_img)

     print('Predicted value is '+ str(np.argmax(ans)))
     img = img.reshape((28,28))
     plt.imshow(img)

     plt.show()
```

Predicted value is 5