

# Session 06

## Linear Regression

**Machine Learning | Hobot | Zahra Amini**

Telegram: @zahraamini\_ai & Instagram:@zahraamini\_ai & LinkedIn: @zahraamini-ai

<https://zil.ink/zahraamini>

# یادگیری ماشین یا Machine Learning

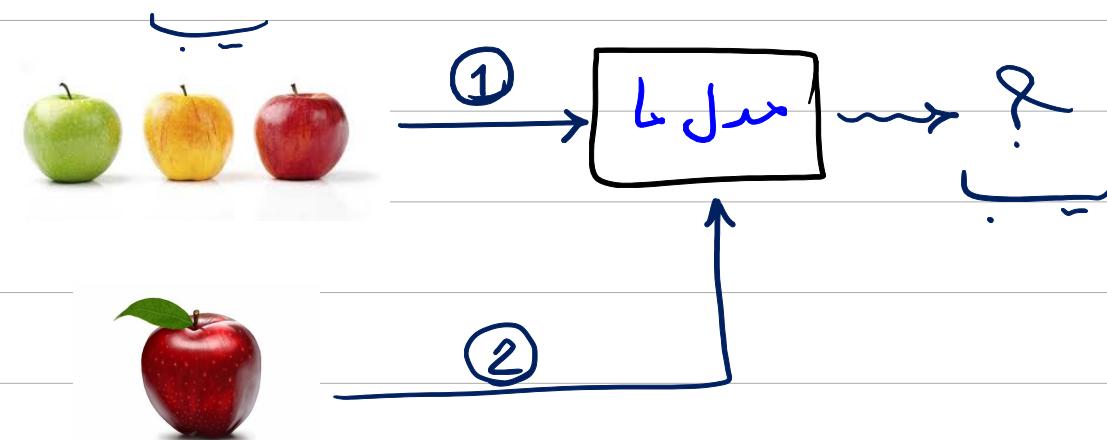
چیست؟ یکی از زیر گروههای هوش مصنوعی است که به کامپیوتر (ستم) این امکان را دهد که به صورت ML

Arthur Samuel

خود کار یادگیرید و پیشرفت کند بعده اینکه به برنامه نویسی صریحی نیاز داشته باشد.

ML → Supervised

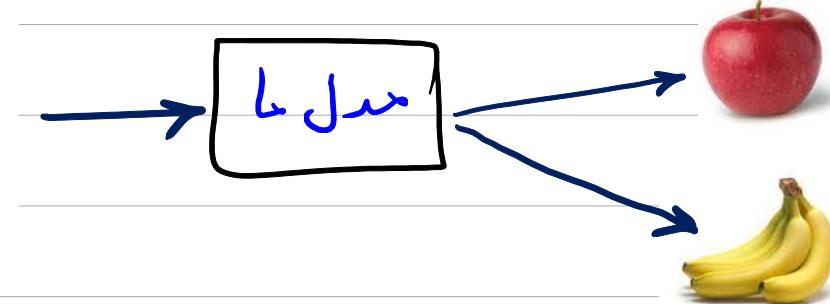
داده‌های label دارند. → نظارتی



Unsupervised

غیر نظارتی

داده‌های بدون label



# Supervised Learning

$X \rightarrow Y$   
input      output  
              label

Email Spam (0,1)

House Price

Supervised

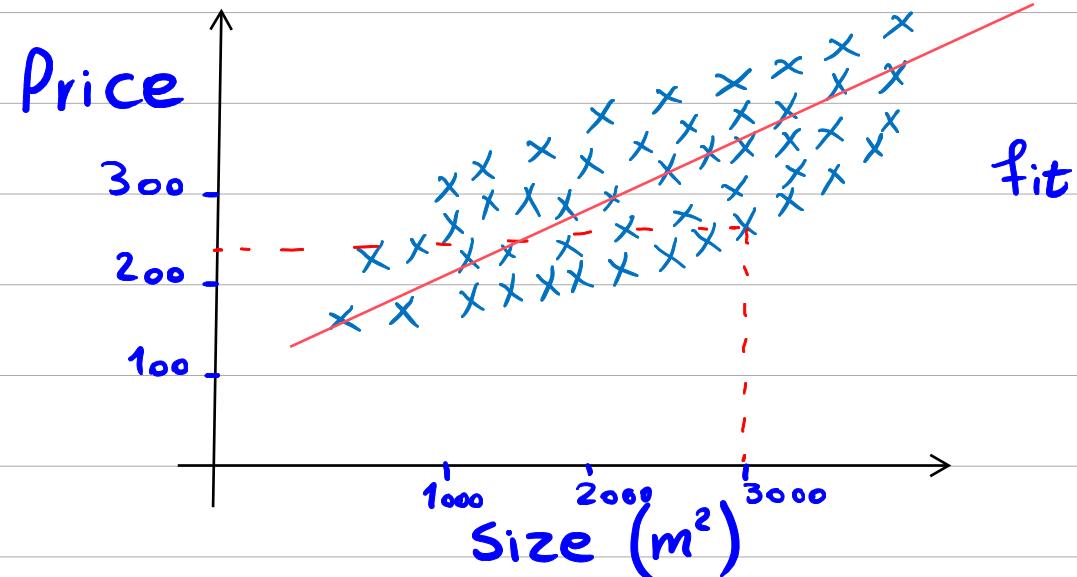
Regression

Classification

Linear Regression  $\rightarrow$  Housing Price Prediction

خواهیم با توجه به مترادر خانه قیمت را پیش بینی کنیم:

Size ( $m^2$ )	Price
180	1000
200	1200
300	2000
:	:



نکته هم: در خروجی ممکن است نهایت عددی توانیم داشته باشیم، بنابراین از رکرسیون برای پیش‌بینی

Continuous

→ feature

→ target

X

y

	Size(m <sup>2</sup> )	Price
(1)	2104	460
(2)	1416	232
(3)	1534	315
(4)	852	178
:	:	:

Training Set

X → مدل → y

متغیرهای پیوسته می‌توان استفاده کرد.

Notation:

X : Input

y: output

m: # training examples 4

n: # features 1

(X, y)

$$x^{(1)} = 2104 \quad y^{(1)} = 460$$

$$(x^{(1)}, y^{(1)}) = (2104, 460)$$

→ (x<sup>(i)</sup>, y<sup>(i)</sup>) : i<sup>th</sup> training example

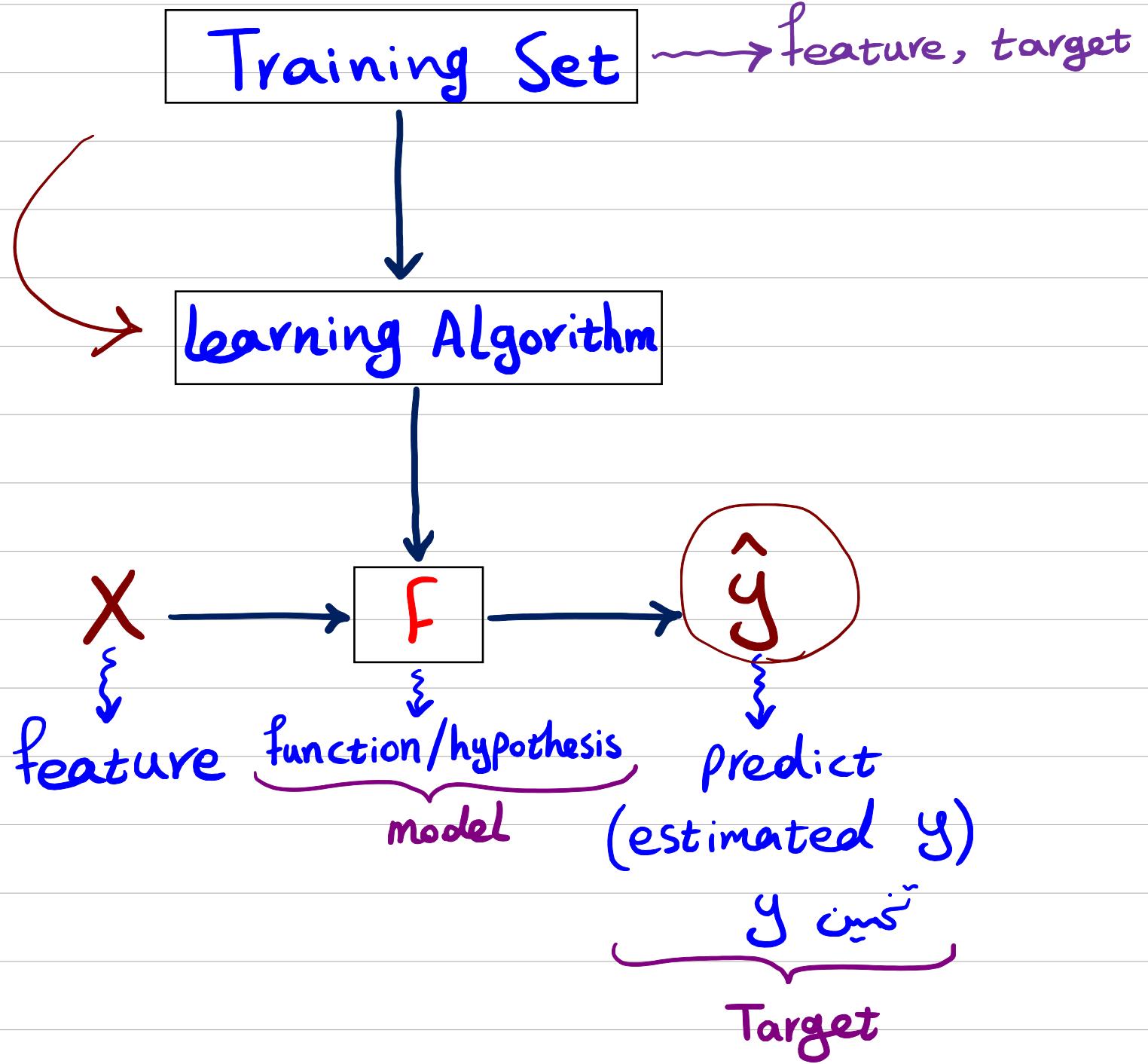
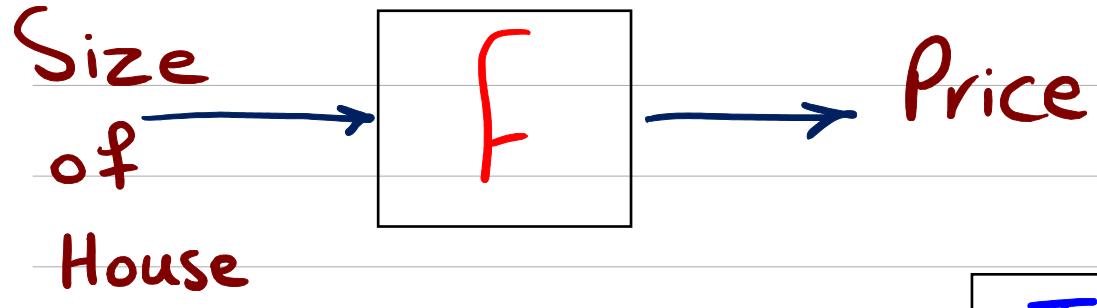
# Test Set

X

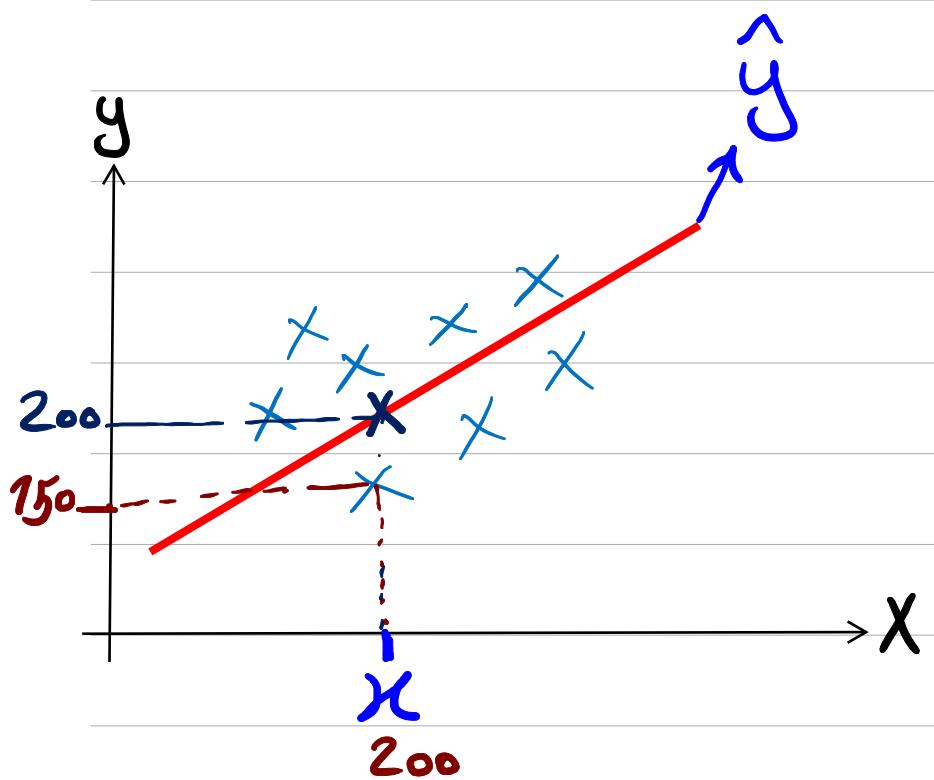
Size(m <sup>2</sup> )
350
2100
833
120
:

ŷ

d. d. d. d. d.



?  $\leftarrow$   $f : ?$

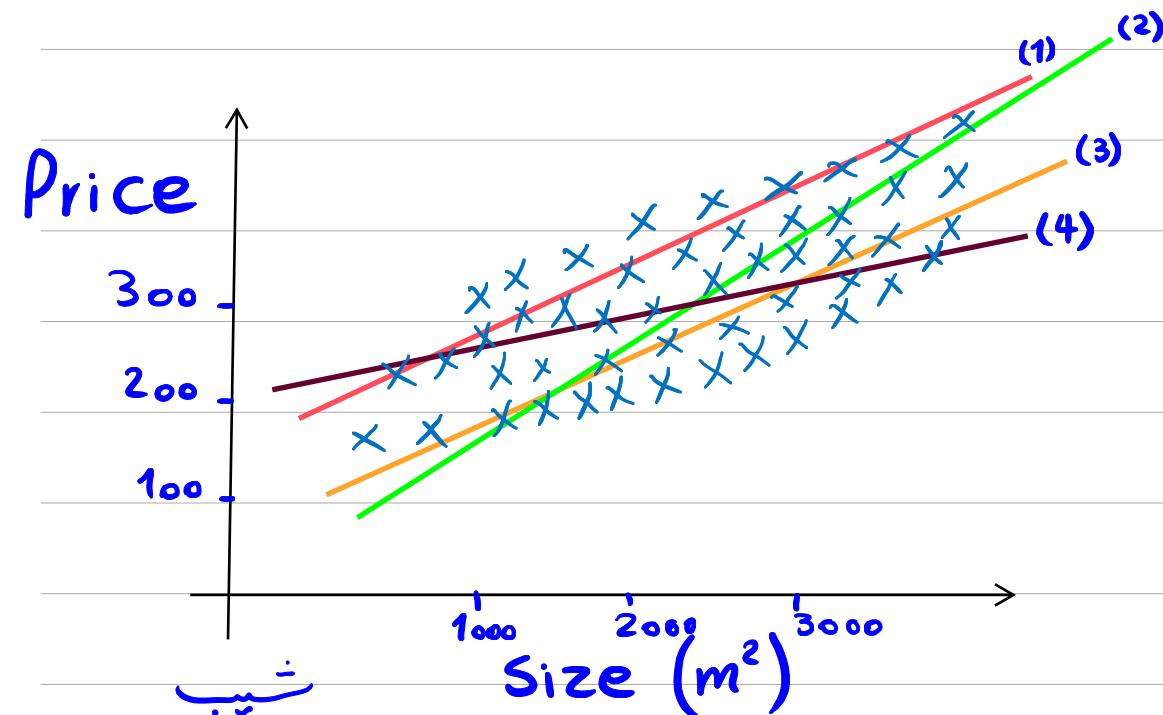


$$y = ax + b \rightsquigarrow \text{Line}$$

$$f_{w,b}(x) = WX + b \rightsquigarrow \text{Linear function}$$

$$f(x) = WX + b$$

فہرستِ کارڈ کا حصہ:  $f_w(x)$ :



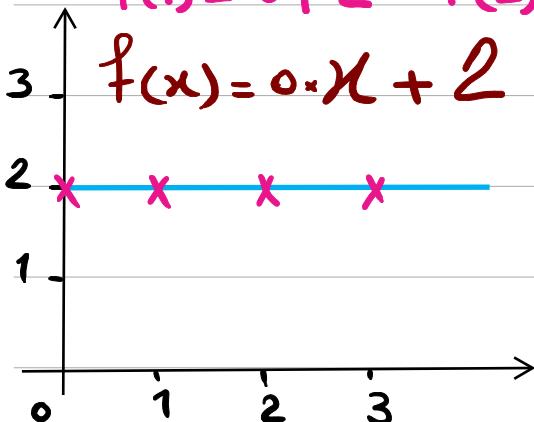
Model:  $f_{w,b}(x) = wx + b$

parameters:  $w, b$

$$f = w \uparrow x + b \rightarrow \text{عرض جبکہ}$$

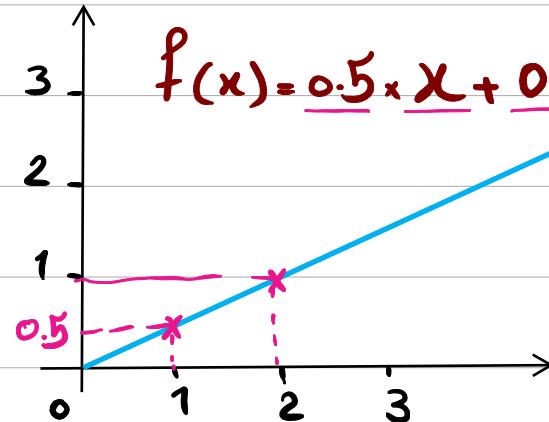
$$f(1) = 0 + 2 \quad f(2) = 2$$

$$f(x) = 0 \cdot x + 2$$

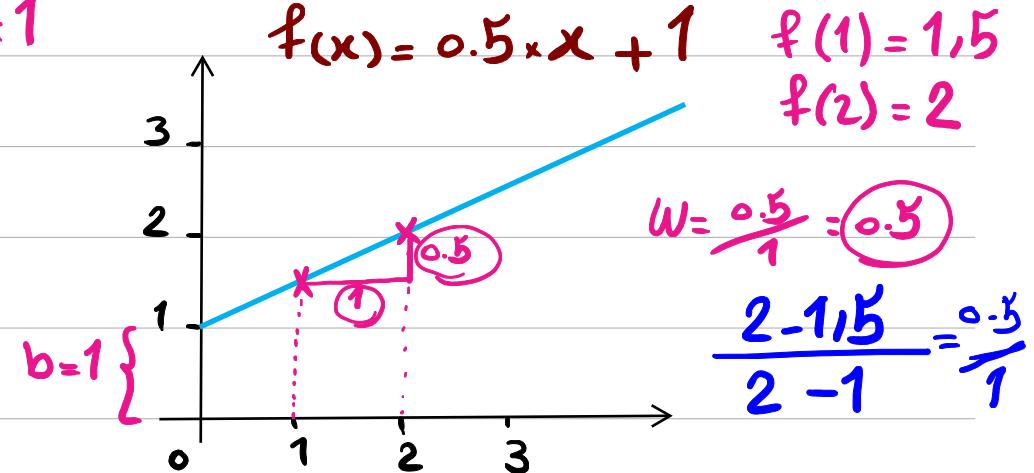


$$f(1) = 0.5 \quad f(2) = 1$$

$$f(x) = 0.5 \times x + 0$$



: Cost Function



$$w=0, b=2$$

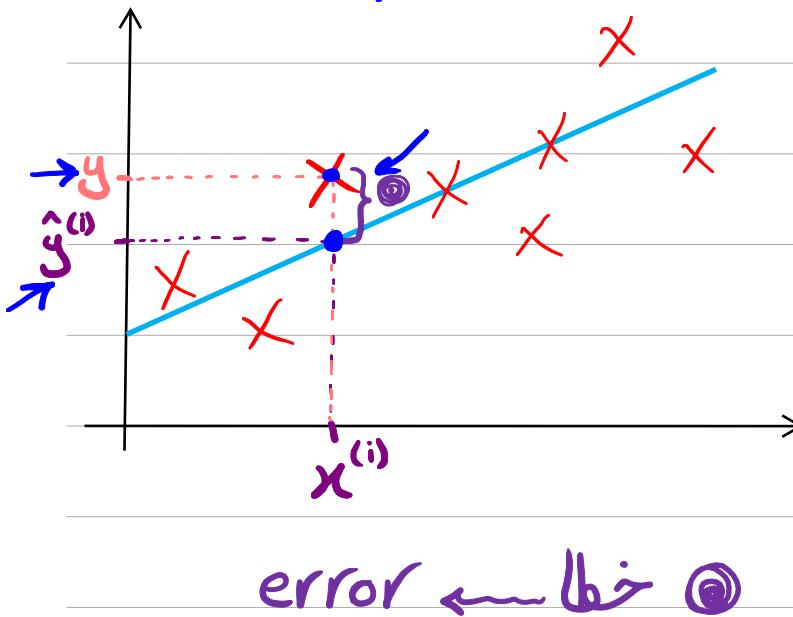
$\hookrightarrow$  Slope

ثیب

$$w=0.5, b=0$$

$$w=0.5, b=1$$

چه  $w$  و  $b$  خوب هستند؟



چه  $w$  و  $b$  خوب هستند؟

و  $w$  و  $b$  خوب است که  $\hat{y}$  به مابعد  $y$  نزدیکتر باشد.

$$\hat{y}^{(i)} \approx y^{(i)} \rightarrow (x^{(i)}, y^{(i)})$$

Cost function:

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m \underbrace{(\hat{y}^{(i)} - y^{(i)})^2}_{\text{error}}$$

$m$ : # training examples

یکی از معروف ترین  $J(w, b)$  cost function های خطا است.

Mean Square Error

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2$$

خط طبقه بارهای پارامترهای  $w, b$  برای مدل علاوه بر مدل:

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)})^2$$

$$\hat{y} = f(x) = wx + b$$

Ex:

Model  $\rightarrow f_{w,b}(x)$

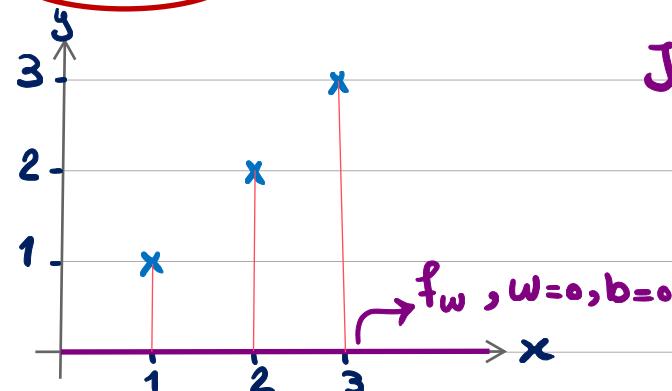
Parameters:  $w, b = 0$

$x$	$y$
1	1
2	2
3	3

Cost function  $\rightarrow J(w) = \frac{1}{2m} \sum_{i=1}^m (f_w(x^{(i)}) - y^{(i)})^2$

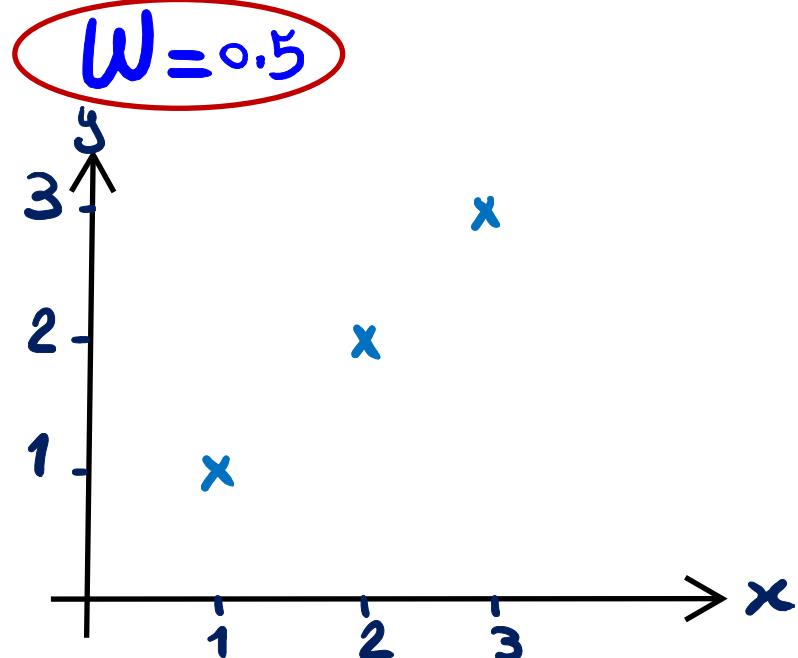
Goal  $\rightarrow$  minimize  $J(w, b)$

$w = 0$



$$J(0) = \frac{1}{2 \times 3} \left[ (0-1)^2 + (0-2)^2 + (0-3)^2 \right] = \frac{1}{6} [1+4+9] = \frac{13}{6} = 2.3$$

$$\rightarrow \hat{y} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$



$(x_i, y)$

$m = 3$

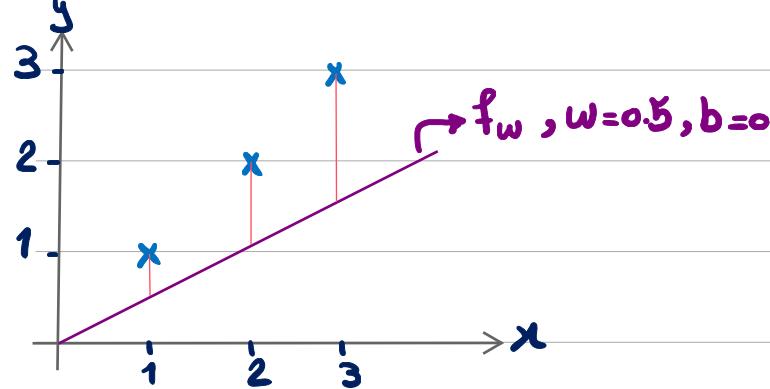
$x$	$y$
1	1
2	2
3	3

$$J(w) = \frac{1}{2m} \sum_{i=1}^m (f_w(x^{(i)}) - y^{(i)})^2$$

$$f_w(x) = wx + b$$

$J(0.5) = ?$

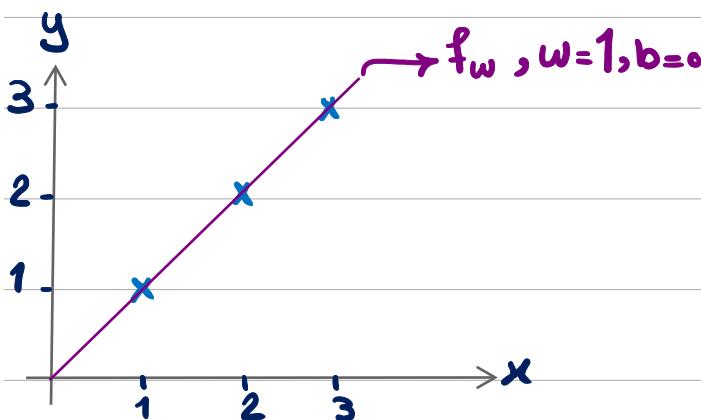
$$W=0.5$$



$$J(0.5) = \frac{1}{6} \left[ (0.5-1)^2 + (1-2)^2 + (1.5-3)^2 \right] = \frac{1}{6} (3.5) = 0.58$$

$$\hat{y} = \begin{bmatrix} 0.5 \\ 1 \\ 1.5 \end{bmatrix}$$

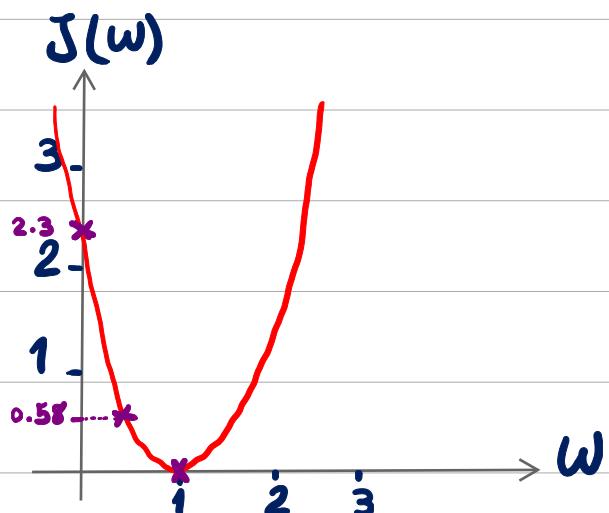
$$W=1$$



$$J(1) = \frac{1}{6} \left[ (1-1)^2 + (2-2)^2 + (3-3)^2 \right] = 0$$

حاله دنبال کنترین خطابویم. کدام  $w$  کنترین خطارا به مایدحد؟

$$\hat{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

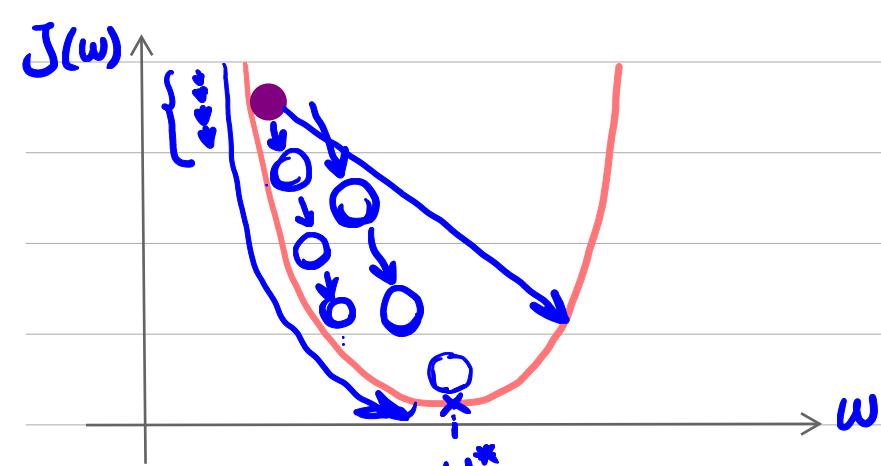


خب حالی خواهیم بود مسیر الگوریتمی که تابع  $J$  را به حداقل برساند.

function  $\rightarrow J(w, b)$

از GD برای به حداقل رساندن هر تابعی که توان استفاده کرد.

Goal  $\rightarrow \min J(w, b)$



$$\rightarrow w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

$$f_{w,b}(x) = wx + b \xrightarrow{\frac{\partial}{\partial w}} x \quad * f^2 = 2f'f$$

$$\frac{\partial}{\partial w} J(w, b) = \frac{\partial}{\partial w} \frac{1}{2m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)})^2 = \frac{\partial}{\partial w} \frac{1}{2m} \sum_{i=1}^m (wx^{(i)} + b - y^{(i)})^2$$

$$= \frac{1}{2m} \sum_{i=1}^m (wx^{(i)} + b - y^{(i)}) \cancel{\times 2} x^{(i)} = \frac{1}{m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)}) x^{(i)}$$

$$f_{w,b}(x) = wx + b \xrightarrow{\frac{\partial}{\partial b}} 1$$

$$\begin{aligned}\frac{\partial}{\partial b} J(w, b) &= \frac{\partial}{\partial b} \frac{1}{2m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)})^2 = \frac{\partial}{\partial b} \frac{1}{2m} \sum_{i=1}^m (wx^{(i)} + b - y^{(i)})^2 \\ &= \frac{1}{2m} \sum_{i=1}^m (wx^{(i)} + b - y^{(i)}) \times 2 = \frac{1}{m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)})\end{aligned}$$

GD Algorithm:

$$w=0, b=0$$

repeat until convergence

{

$$w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

}

$x$	$y$	$\hat{y}$
1	1	1
2	3	1
3	3	1

برای دفعه اول رگرسیون خطی را بعداز؟ بار آبدیت بدست آورید.  
 $(\alpha=0.1, b=1, w=0)$

$$w := w - \alpha \frac{\partial}{\partial w} J(w, b) \rightarrow w_{\text{new}} = 0 - \left(\frac{1}{10}\right) \times \left(-\frac{10}{3}\right) = 0 + \frac{1}{3} = 0.33$$

$$\cancel{\frac{1}{m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)}) x^{(i)}} = \frac{1}{3} \times \left[ \cancel{(1-1)} + \cancel{(1-3)} + \cancel{(1-3)} \right] = -\frac{10}{3}$$

$$b := b - \alpha \frac{\partial}{\partial b} J(w, b) \rightarrow b_{\text{new}} = 1 - \underline{\underline{\left(\frac{1}{10}\right) \times \left(-\frac{4}{3}\right)}} = 1.13$$

$$\cancel{\frac{1}{m} \sum_{i=1}^m (f_{w,b}(x^{(i)}) - y^{(i)})} = \frac{1}{3} \times \left[ \cancel{(1-1)} + \cancel{(1-3)} + \cancel{(1-3)} \right] = -\frac{4}{3}$$

$$\hookrightarrow w_{\text{new}} = 0.33$$

$$b_{\text{new}} = 1.13$$

$$f_{w,b}(x) = 0.33x + 1.13$$