

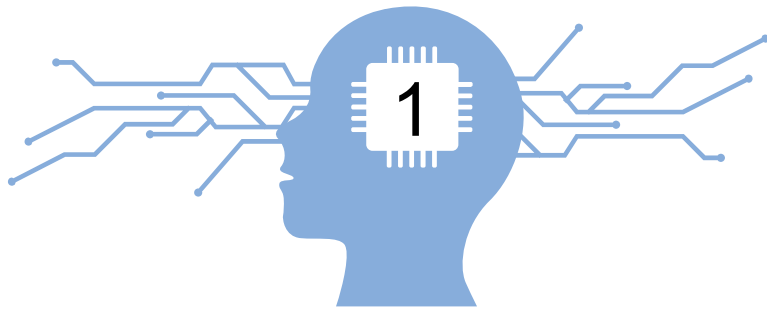


Kecerdasan Buatan (Artificial Intelligence)

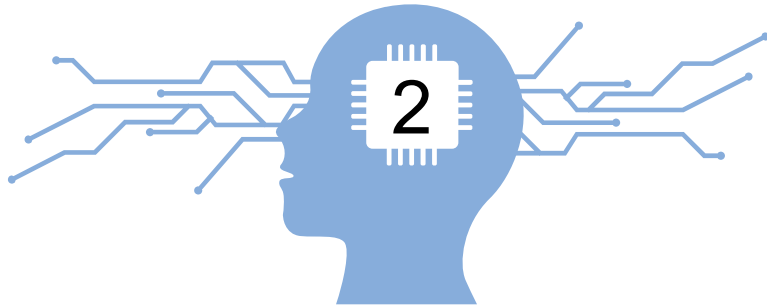
Regression

Chandra Prasetyo Utomo

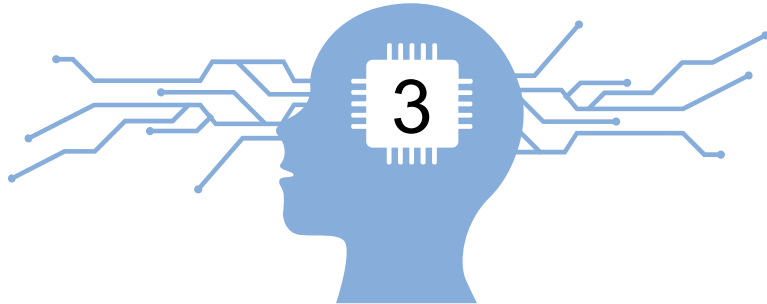
Learning Objectives



Understand supervised learning training and testing procedure.



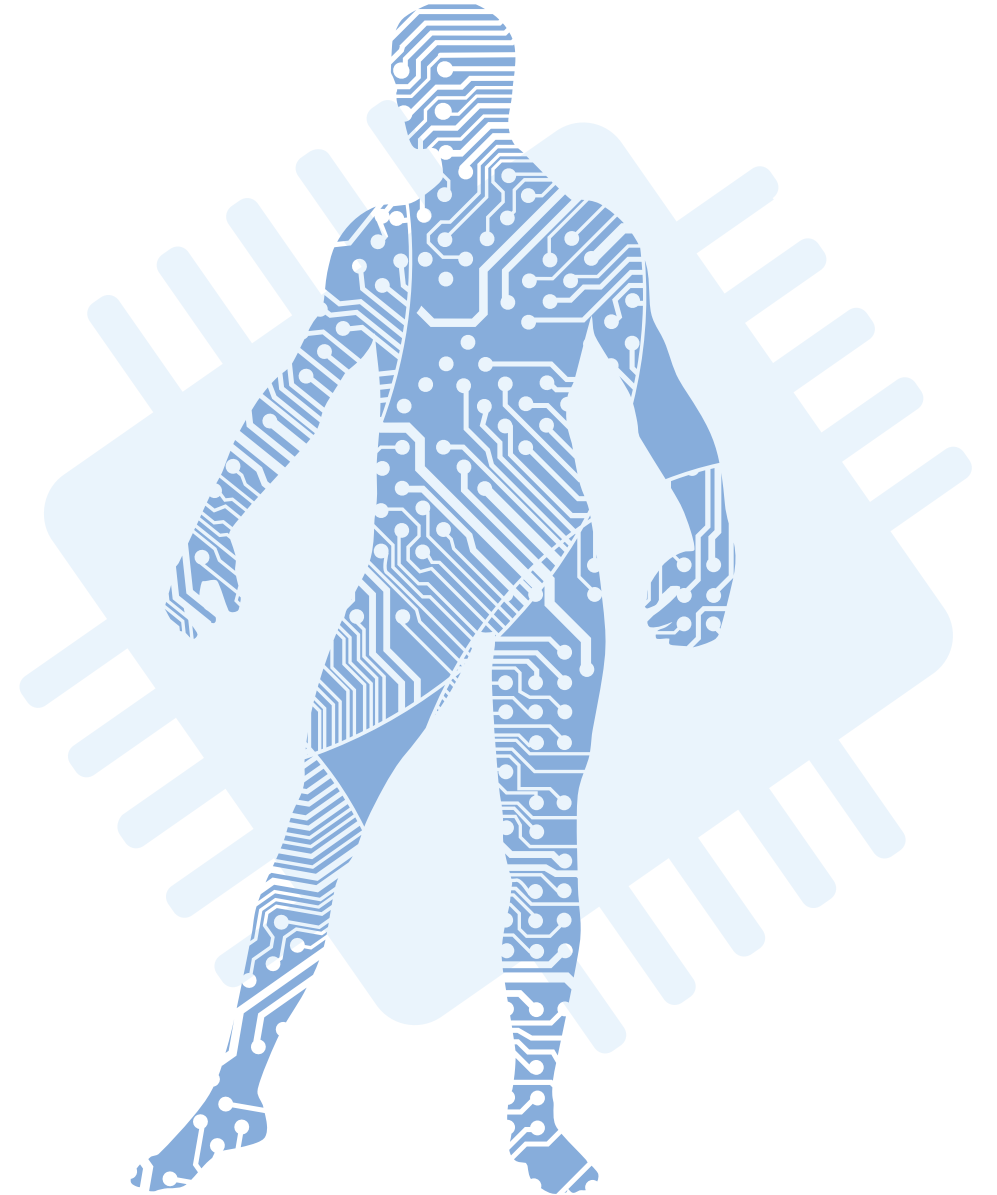
Understand model representation in linear regression algorithm.

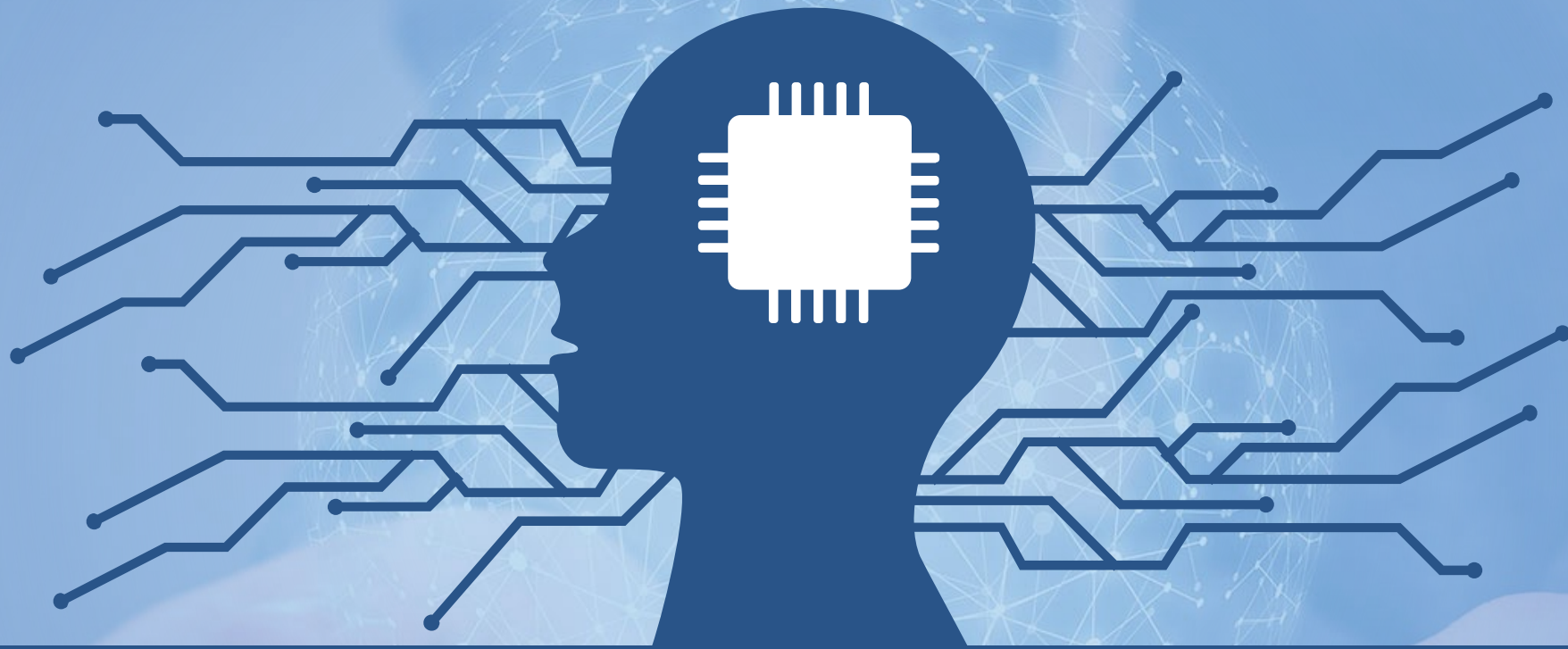


Understand model learning using gradient descent for linear regression.

Lecture Outline

- 1 Supervised Learning**
- 2 Model Representation**
- 3 Model Learning**





Section 1

Supervised Learning

House Price Prediction in East Bekasi

Rumah Dijual di Bekasi Timur, i x +

rumah.com/proerti-dijual?area_code=IDJB04007&autocomplete=%7B%22objectid%3A%22IDJB04007%2C%22objectType%3A%22AREA%2C%22propertie...

820 Rumah Dijual, Diantara Rp 500 Jt dan Rp 2 M, di Bekasi Timur, Bekasi Ditemukan

Log In atau Daftar 0 ♥

Jawa Barat, Bekasi, Bekasi Timur

Beli

Rumah

Rp 500,000,000 - 2,00...

Filter

Cari

FEATURED

13

Didaftarkan oleh Anto.
"Dijual rumah gratis dp"

Hubungi

Rumah Mewah Gratisan DP dan Biaya Pros...

Jl. Wibawa Mukti II, Jatiasih, Bekasi Timur, Bekasi, Jawa ...

Rp 700 jt

3 2 82 m² · Rp 7.777.778 per m²

Rumah

SHM - Sertifikat Hak Milik

3 minggu

8

PanggilanVideo

Rumah 2 lantai harga 1 lantai

Jl.bengkong, Bekasi Timur, Bekasi, Jawa Barat

Rp 600 jt

2 2 51 m² · Rp 9.523.810 per m²

Rumah

SHM - Sertifikat Hak Milik

58 menit

Didaftarkan oleh Siti Rosidah
"Rumah 2 lantai harga 1 lantai"

Hubungi

Perumahan Syariah Di Kota Bekasi

14

1

Grand Al Ihsan Premier

Mustika Sari, Bekasi Timur, Bekasi, Jawa Barat

Rp 755 jt

3 2 41 m² · Rp 10.486.111 per m²

Rumah

SHM - Sertifikat Hak Milik

2 jam

Didaftarkan oleh Septi Handayani
"Dijual rumah 2 km ke tol Bekasi Timur cuma 755 jt"

Hubungi

Hendrik Pangharapan

F1 Propertiindo

+62 813 8212

Petunjuk Agen >

BUMI

BNI

BNI Griya Expo Online

CUKUP BAYAR LANGSUNG PINDAH

BUNGA RUMAH

Harga di [bungrumah.com](#)

17 Juli - 17 Agustus 2024

Properti Baru Di Area Ini

TERA DAMAI (CLUSTER ASTAMA)

Damai Putra Group Jalan Turi Keluaran Sriamur

Royal Oak Town House

Justin Leon Setiawan Jl. Pertamina B No.9-10, RT.002/RW.006

SAMIRA REGENCY BEKASI

PT TRIYASA PROPERTINDO Jl. Gondang, RT.003/RW.005, Cimuning,

Size (m ²)	Bed	Bath	...	Price (IDR million)
82	3	2	...	700
51	2	2	...	600
41	3	2	...	755
...

Size (m ²)	Bed	Bath	...	Price (IDR million)
98	4	3	...	???
45	3	1	...	???
36	2	1	...	???
...

House Price Prediction: Simplified

Training Data

x	y
Size (m ²)	Price (IDR million)
82	700
51	600
41	755
...	...

Testing Data

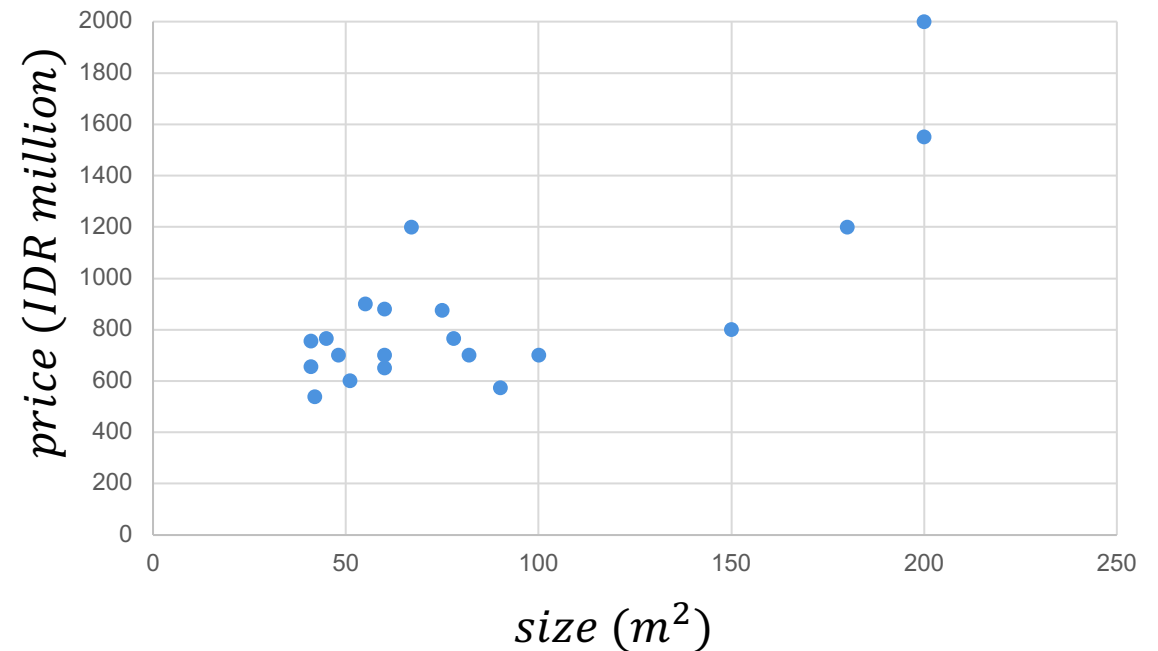
x	y
Size (m ²)	Price (IDR million)
98	???
45	???
36	???
...	...

$x = \text{size} ; y = \text{price}$

$$h(x) = y$$

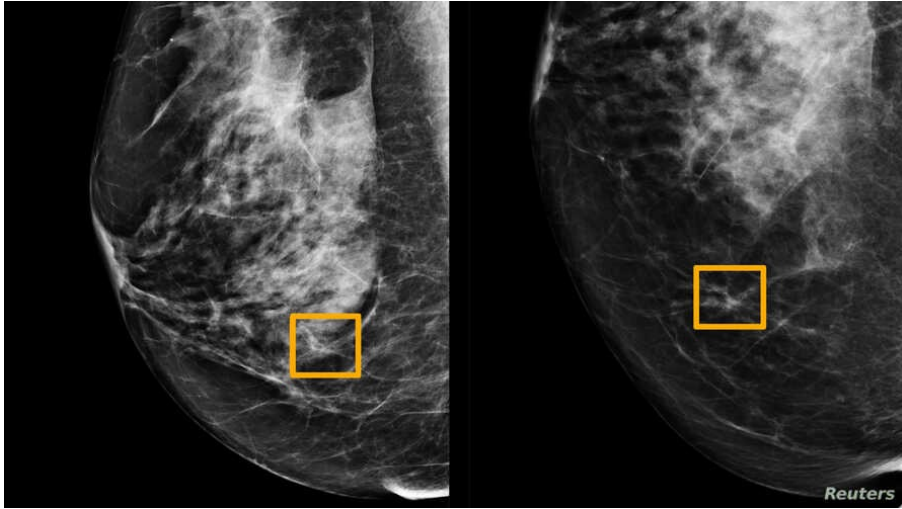
Because y is a **continuous variable**,
then this will be a **REGRESSION** problem.

House Price in East Bekasi



if y was a **discrete variable (class)**,
then this would be a **CLASSIFICATION** problem.

Breast Cancer Diagnostic



Benign (B): not harmful
Malignant (M): deadly

Clump Thickness	Cell Size	Cell Shape	...	Diagnostic
5	4	4	...	B
8	10	10	...	M
4	1	1	...	B
...

Clump Thickness	Cell Size	Cell Shape	...	Diagnostic
6	8	8	...	?
3	2	1	...	?
7	4	6	...	?
...

Breast Cancer Diagnostic: Simplified

Training Data

x	y
Cell Size	Diagnostic
4	B
10	M
1	B
...	...

Testing Data

x	y
Cell Size	Diagnostic
8	?
2	?
4	?
...	...

$x = \text{cell size}$
 $y = \text{diagnostic}$

$$h(x) = y$$

*Because y is a **discrete variable (class)**,
then this is a **CLASSIFICATION** problem.*

Supervised Learning Procedure

Training Data

Size (m ²)	Price (IDR million)
82	700
51	600
41	755
...	...

Training Data



Machine Learning Algorithm



Trained Model

Testing Data

Size (m ²)	Price (IDR million)
98	???
45	???
36	???
...	...

Testing Data

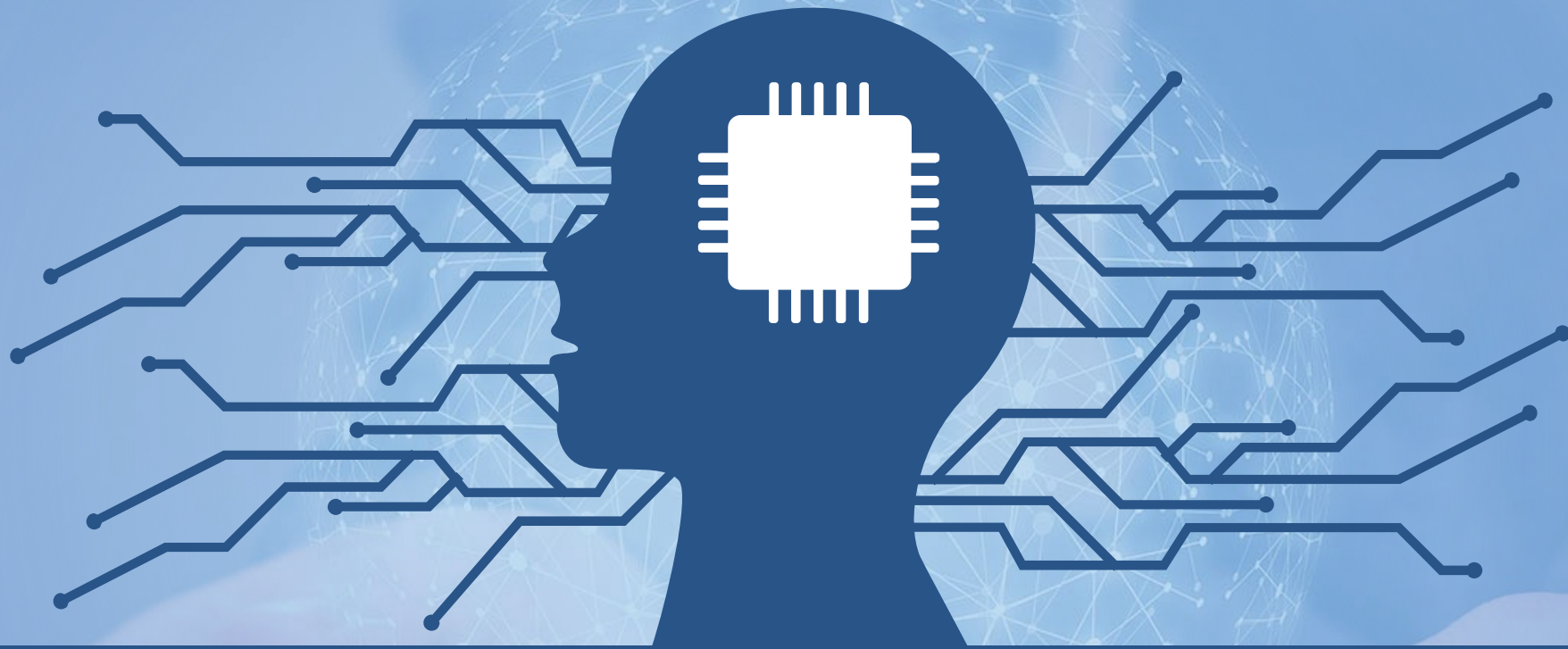


Prediction

x

$h(x) = y$

y



Section 2

Model Representation

Notation

*Training Data of Housing Prices
(East Bekasi)*

	x	y
	size in m ²	Price (IDR) in million
1	82	700
2	51	600
3	41	755

m	67	1200

x = features / input variable

y = target / output variable

m = number of training examples

(x, y) = a single training example

$(x^{(i)}, y^{(i)})$ = i^{th} training example

$$x^{(1)} = 82$$

$$x^{(2)} = 51$$

$$y^{(2)} = 600$$

$$(x^{(3)}, y^{(3)}) = (41, 755)$$

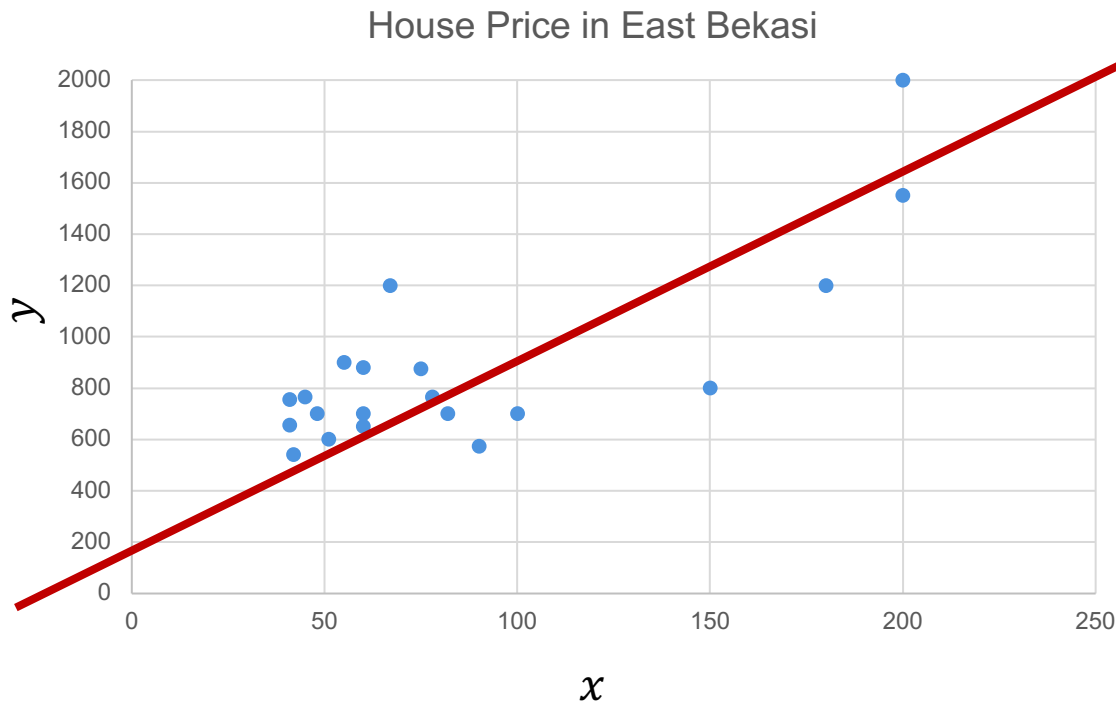
Linear Model

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

h : a linear model,
maps from x to y

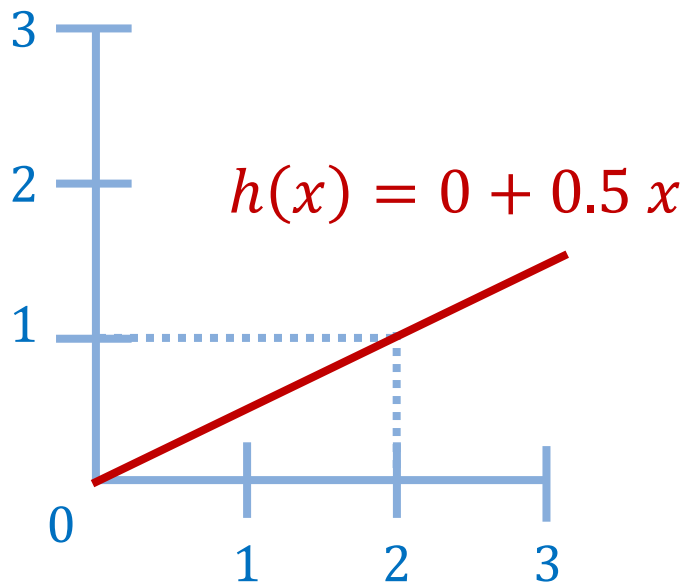
θ_0 , θ_1 : parameters

x : input

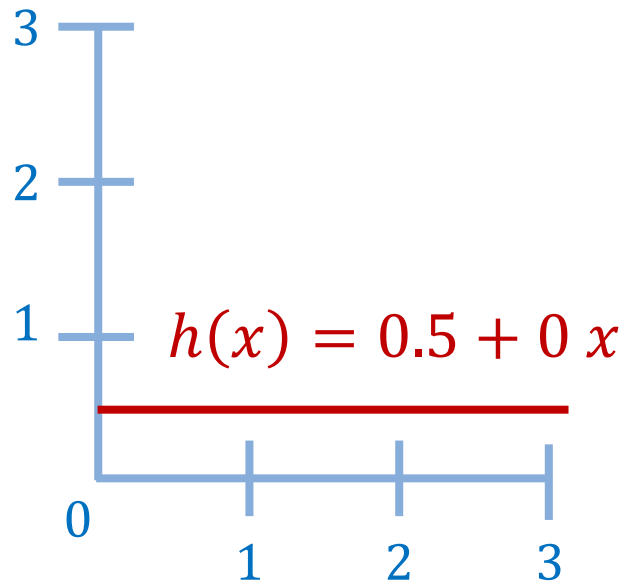


Hypothesis

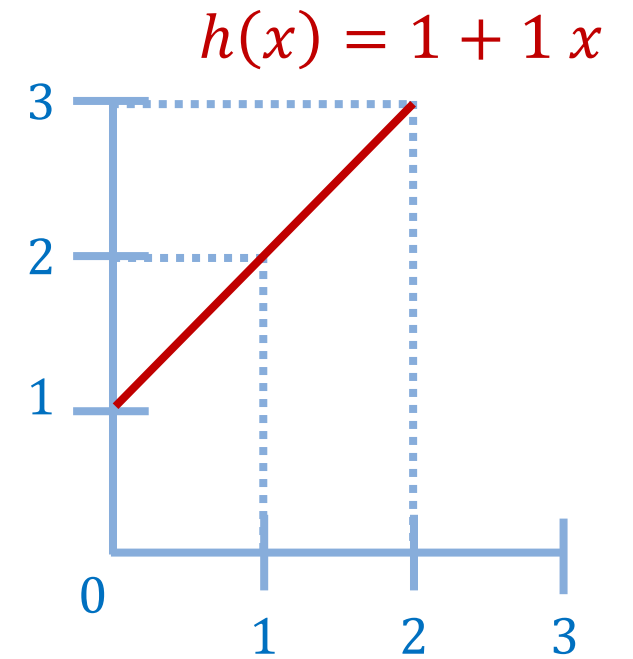
$$h_{\theta}(x) = \theta_0 + \theta_1 x$$



$$\begin{aligned}\theta_0 &= 0 \\ \theta_1 &= 0.5\end{aligned}$$

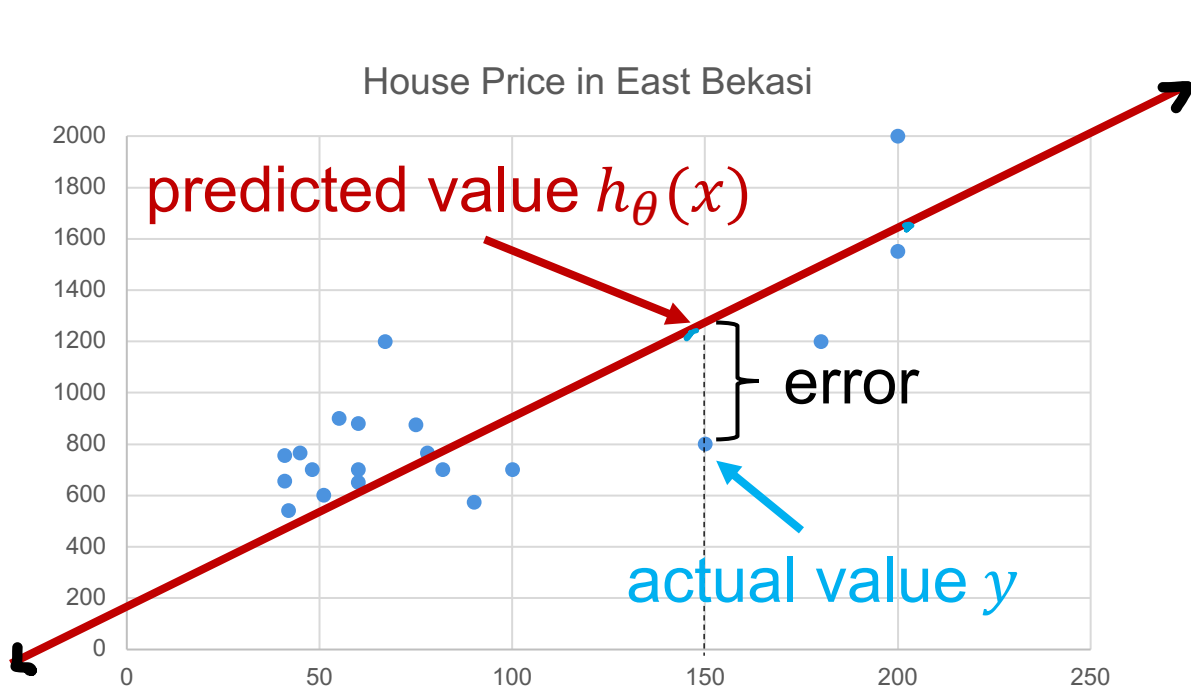


$$\begin{aligned}\theta_0 &= 0.5 \\ \theta_1 &= 0\end{aligned}$$



$$\begin{aligned}\theta_0 &= 1 \\ \theta_1 &= 1\end{aligned}$$

Mean Squared Error (MSE)



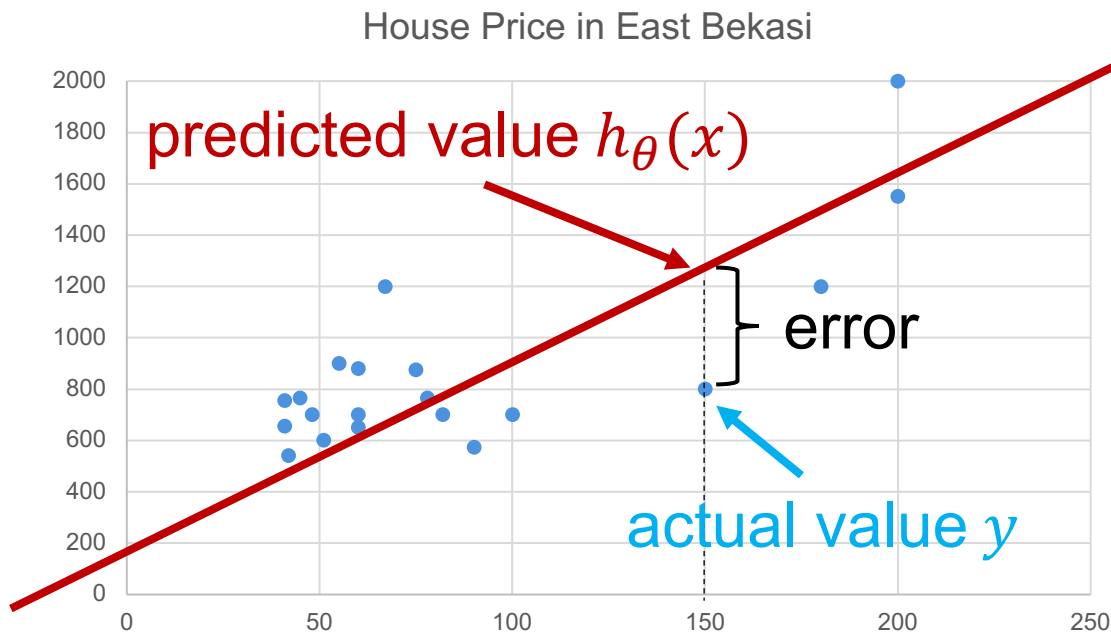
$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

$$\begin{aligned}\text{Error} &= \text{predicted} - \text{actual} \\ &= h_{\theta}(x) - y\end{aligned}$$

$$\begin{aligned}\text{Squared Error} &= \text{Error}^2 \\ &= (h_{\theta}(x) - y)^2\end{aligned}$$

$$\begin{aligned}\text{MSE} &= \frac{1}{m} \{ (h_{\theta}(x^{(1)}) - y^{(1)})^2 + (h_{\theta}(x^{(2)}) - y^{(2)})^2 + \dots + (h_{\theta}(x^{(m)}) - y^{(m)})^2 \} \\ &= \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2\end{aligned}$$

Cost Function $J(\theta_0, \theta_1)$



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

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

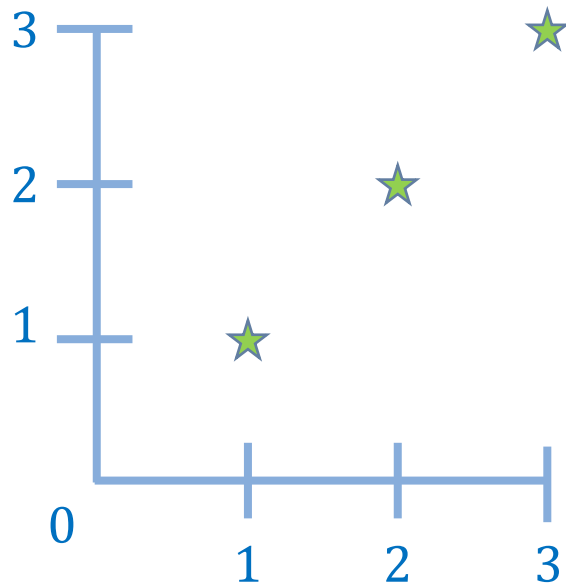
Objective: Find θ_0, θ_1 such that $h_{\theta}(x)$ is close to y for training example (x, y)

$$\underset{\theta_0, \theta_1}{\text{minimize}} J(\theta_0, \theta_1)$$

Cost Function – Example in a Simplified Model

Training Data

x	y
1	1
2	2
3	3



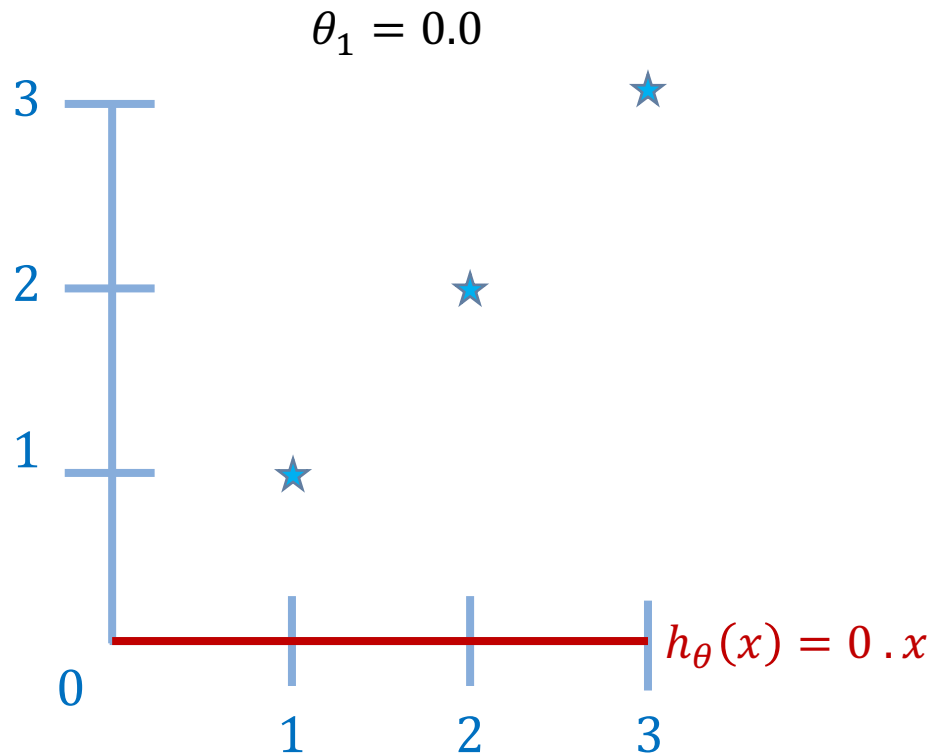
Simplified Model: $h_{\theta}(x) = \theta_1 x$

Parameter: θ_1

Cost Function: $J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$

Objective: minimize $J(\theta_1)$

Cost Function – Example in a Simplified Model



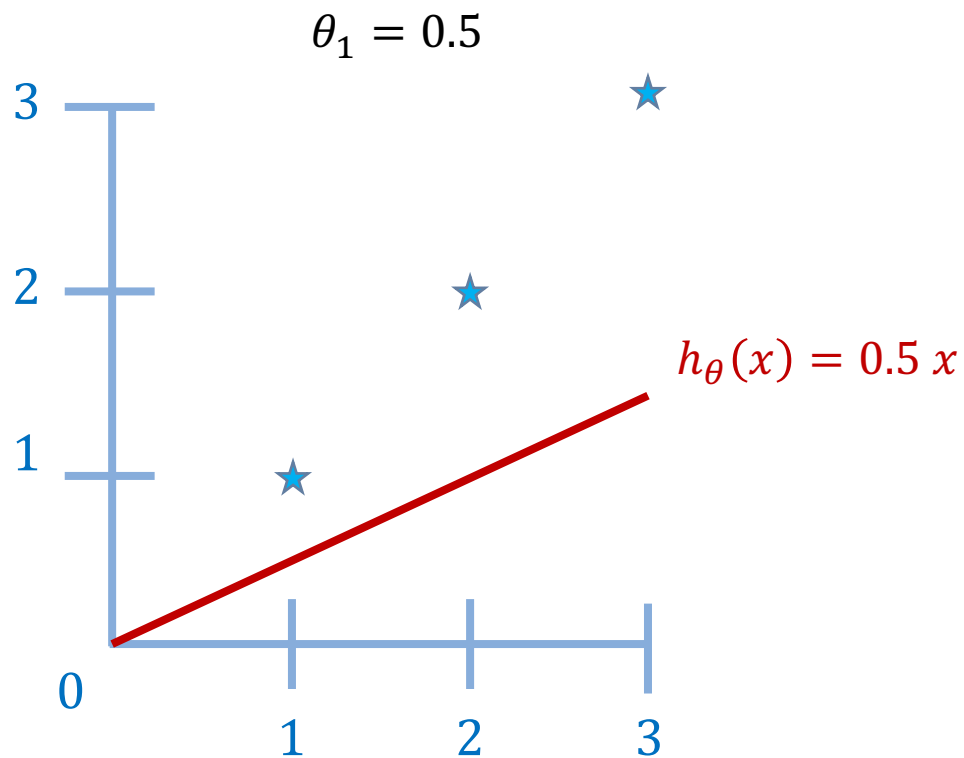
$$\begin{aligned}h_{\theta}(1) &= 0 \cdot \\h_{\theta}(2) &= 0 \cdot \\h_{\theta}(3) &= 0 \cdot\end{aligned}$$

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

$x^{(i)}$	$y^{(i)}$	$h_{\theta}(x^{(i)})$	$h_{\theta}(x^{(i)}) - y^{(i)}$	$(h_{\theta}(x^{(i)}) - y^{(i)})^2$
1	1	0	$0 - 1 = -1$	$(-1)^2 = 1$
2	2	0	$0 - 2 = -2$	$(-2)^2 = 4$
3	3	0	$0 - 3 = -3$	$(-3)^2 = 9$
$\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				$1 + 4 + 9 = 14$
$\frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				$14 / 6 = 2.33$

$$J(0.0) = 2.33$$

Cost Function – Example in a Simplified Model



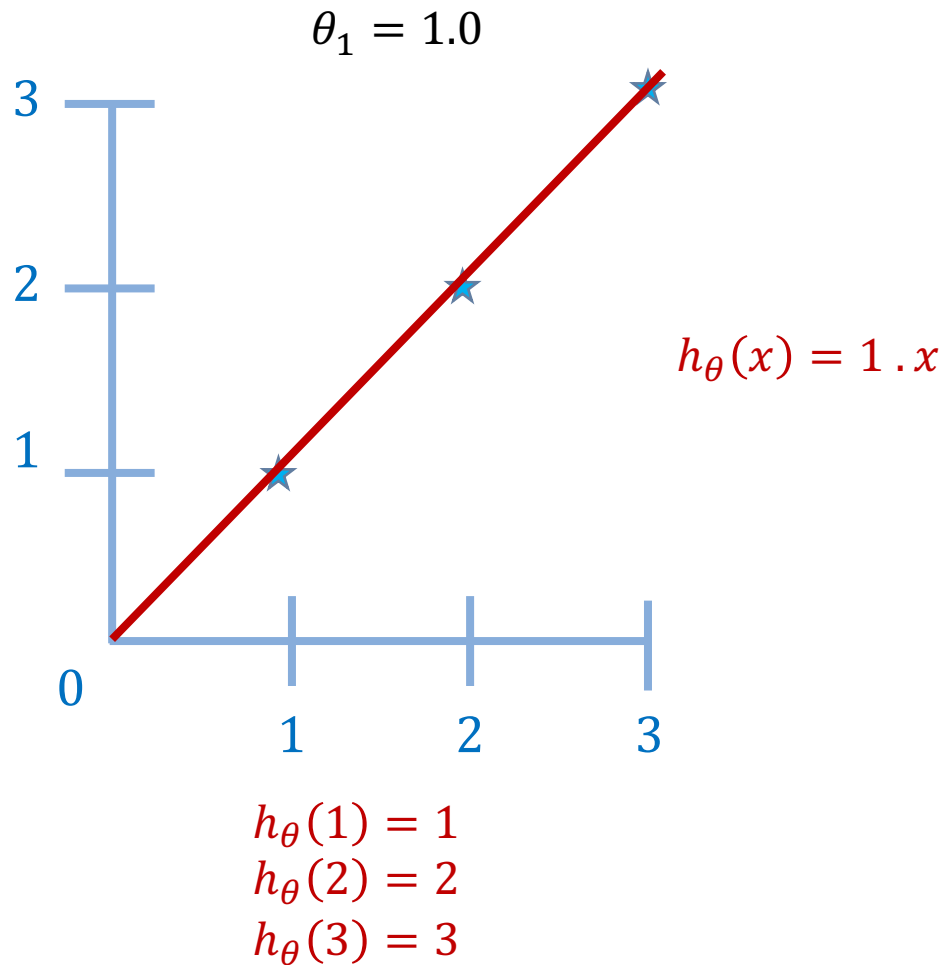
$$\begin{aligned}h_{\theta}(1) &= 0.5 \\h_{\theta}(2) &= 1.0 \\h_{\theta}(3) &= 1.5\end{aligned}$$

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

$x^{(i)}$	$y^{(i)}$	$h_{\theta}(x^{(i)})$	$h_{\theta}(x^{(i)}) - y^{(i)}$	$(h_{\theta}(x^{(i)}) - y^{(i)})^2$
1	1	0.5	$0.5 - 1 = -0.5$	$(-0.5)^2 = 0.25$
2	2	1.0	$1.0 - 2 = -1.0$	$(-1.0)^2 = 1.00$
3	3	1.5	$1.5 - 3 = -1.5$	$(-1.5)^2 = 2.25$
$\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				$0.25 + 1.00 + 2.25 = 3.50$
$\frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				$3.50 / 6 = 0.58$

$$J(0.5) = 0.58$$

Cost Function – Example in a Simplified Model

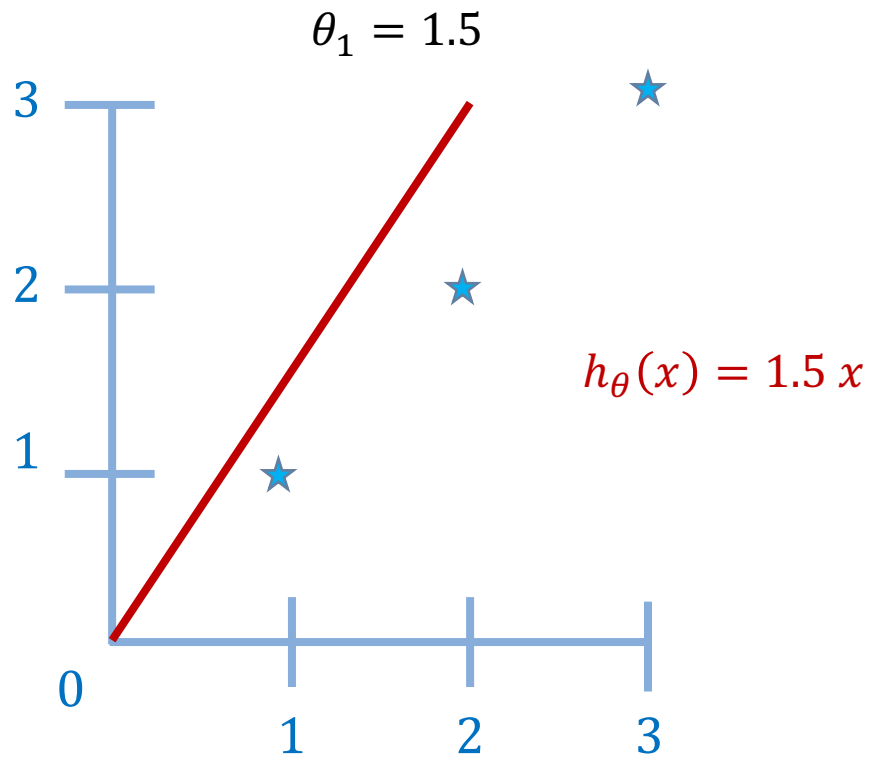


$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

$x^{(i)}$	$y^{(i)}$	$h_{\theta}(x^{(i)})$	$h_{\theta}(x^{(i)}) - y^{(i)}$	$(h_{\theta}(x^{(i)}) - y^{(i)})^2$
1	1	1	0	0
2	2	2	0	0
3	3	3	0	0
$\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				0
$\frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				0

$$J(1.0) = 0$$

Cost Function – Example in a Simplified Model



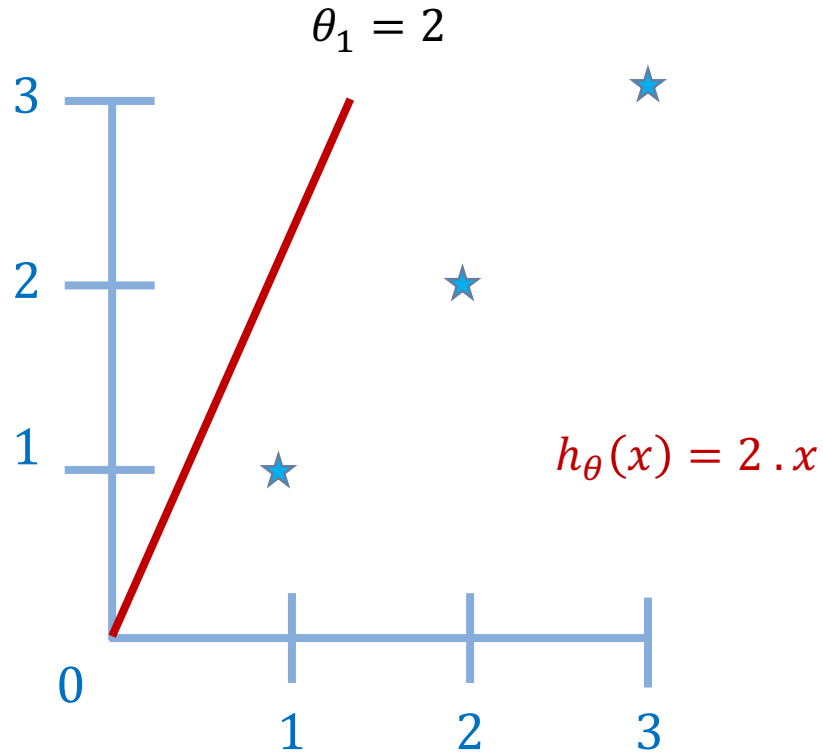
$$\begin{aligned}h_{\theta}(1) &= 1.5 \\h_{\theta}(2) &= 3.0 \\h_{\theta}(3) &= 4.5\end{aligned}$$

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

$x^{(i)}$	$y^{(i)}$	$h_{\theta}(x^{(i)})$	$h_{\theta}(x^{(i)}) - y^{(i)}$	$(h_{\theta}(x^{(i)}) - y^{(i)})^2$
1	1	1.5	0.5	0.25
2	2	3.0	1.0	1.00
3	3	4.5	1.5	2.25
$\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				3.50
$\frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				$3.50 / 6 = 0.58$

$$J(1.5) = 0.58$$

Cost Function – Example in a Simplified Model



$$\begin{aligned}h_{\theta}(1) &= 2 \\h_{\theta}(2) &= 4 \\h_{\theta}(3) &= 6\end{aligned}$$

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

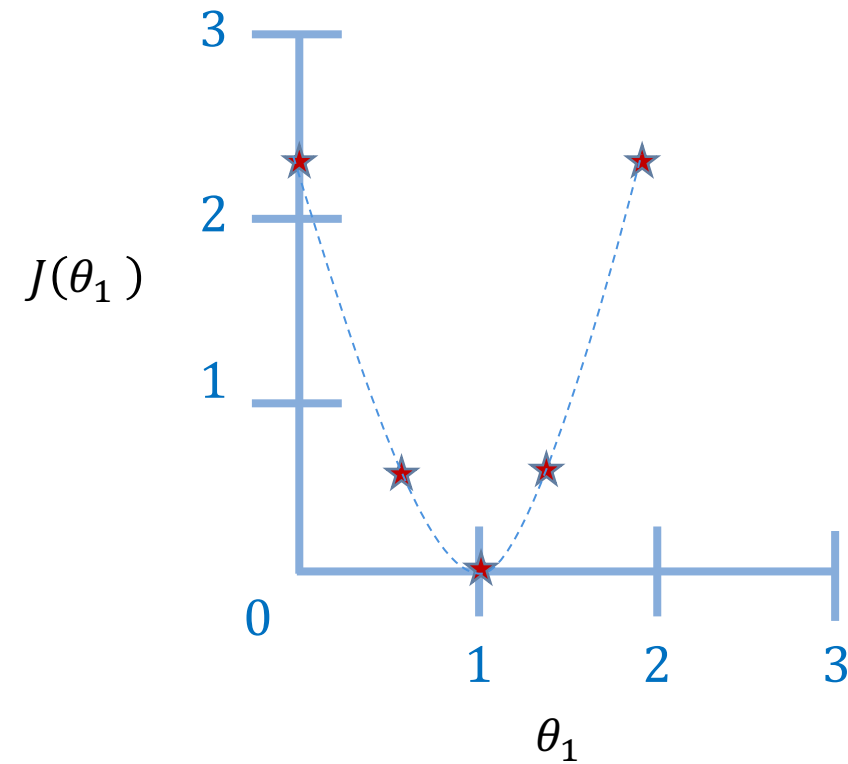
$x^{(i)}$	$y^{(i)}$	$h_{\theta}(x^{(i)})$	$h_{\theta}(x^{(i)}) - y^{(i)}$	$(h_{\theta}(x^{(i)}) - y^{(i)})^2$
1	1	2	1	1
2	2	4	2	4
3	3	6	3	9
$\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				14
$\frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$				14 / 6 = 2.33

$$J(2) = 2.33$$

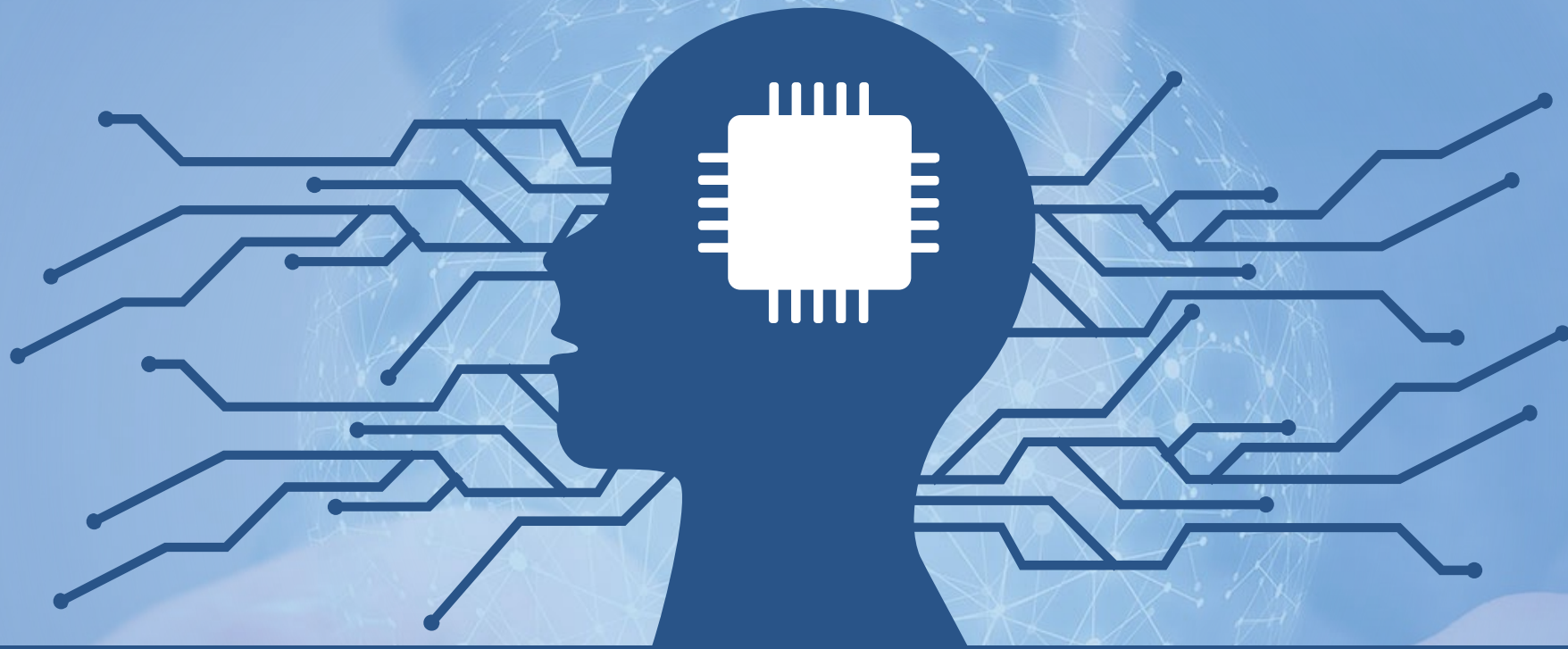
Cost Function – Example in a Simplified Model

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

θ_1	$J(\theta_1)$
0.0	2.33
0.5	0.58
1.0	0.00
1.5	0.58
2.0	2.33



*The closer θ_1 to the optimal value,
the smaller its cost function $J(\theta_1)$*



Section 3

Model Learning

Idea

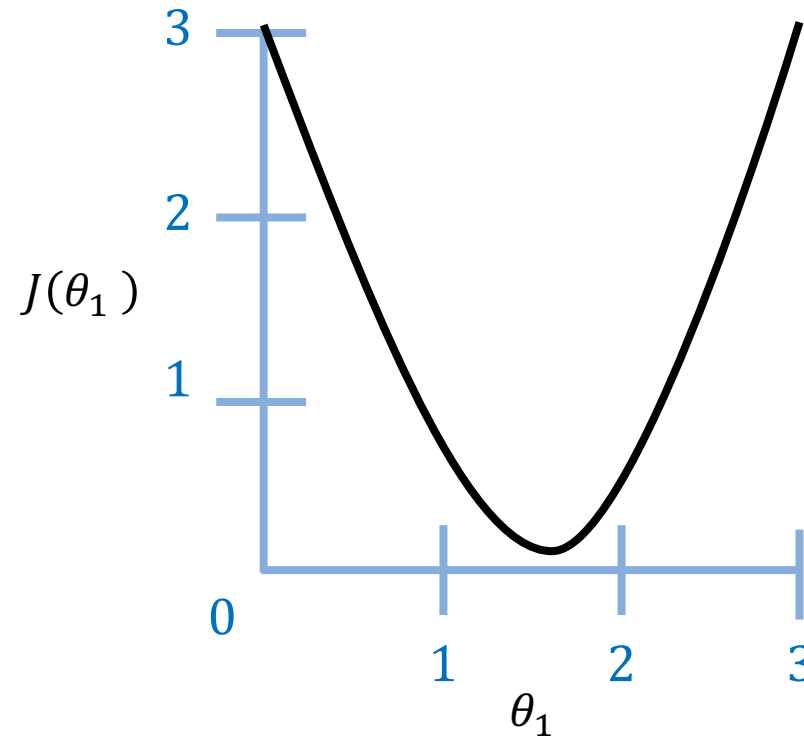
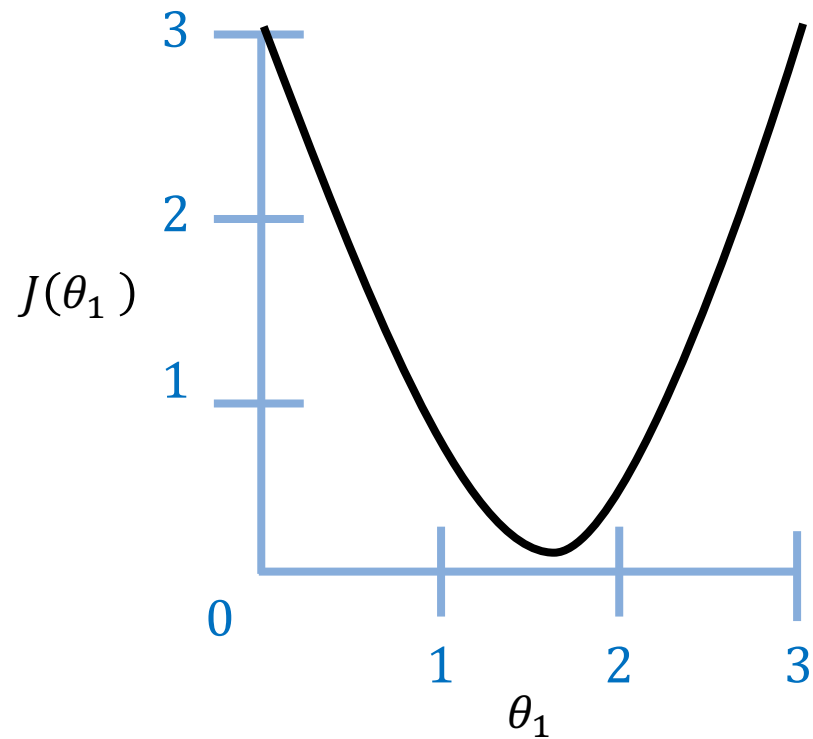
Define a cost function $J(\theta_0, \theta_1)$

Try minimize $J(\theta_0, \theta_1)$
 θ_0, θ_1

Plan:

- Initialize θ_0, θ_1 .
- Iteratively update θ_0, θ_1 to minimize $J(\theta_0, \theta_1)$ until reach minimum point.

Intuition: what we need to “go down”?



Need two things:
1. Direction
2. Movement

Gradient Descent Algorithm

Repeat until convergence {

$$\theta_0 := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$

$$\theta_1 := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$

}

α = “learning rate” (set manually)

$\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$ = “direction” (partial derivative)

Choose Learning Rate α (Manually)

$$\theta_1 := \theta_1 - \alpha \frac{\delta}{\delta \theta_1} J(\theta_0, \theta_1)$$

If α is too small, gradient descent can be slow.

If α is too big, gradient descent can overshoot the minimum (fail to converge).

Compute Partial Derivative

$$\begin{aligned}\frac{\delta}{\delta \theta_j} J(\theta_0, \theta_1) &= \frac{\delta}{\delta \theta_j} \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 \\ &= \frac{\delta}{\delta \theta_j} \frac{1}{2m} \sum_{i=1}^m (\theta_0 + \theta_1 x^{(i)} - y^{(i)})^2\end{aligned}$$

$$\frac{\delta}{\delta \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})$$

$$\frac{\delta}{\delta \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x^{(i)}$$

Gradient Descent for Linear Regression

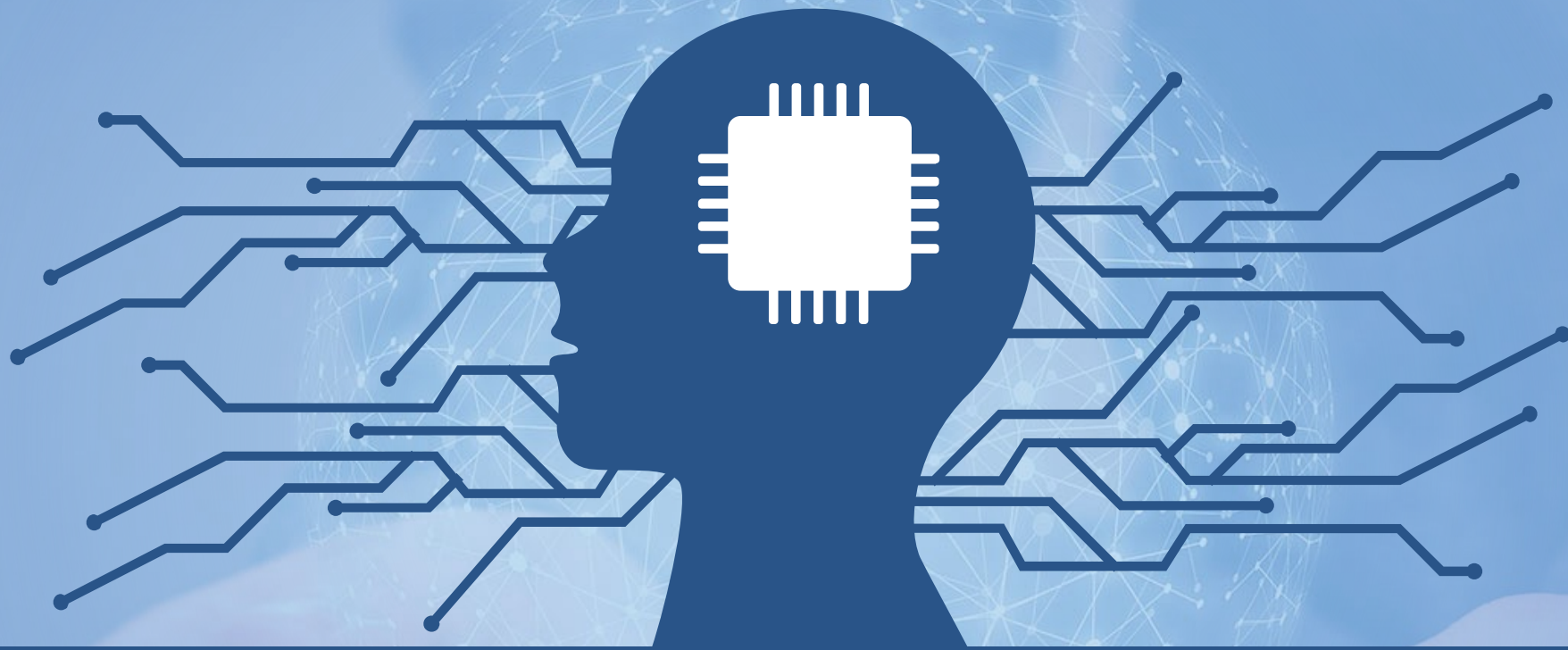
Repeat until convergence {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x^{(i)}$$

}

update
 θ_0 and θ_1
simultaneously



Thank you!

Acknowledgment

Slide Template:

<https://www.free-powerpoint-templates-design.com/>

Slide Creator:

Chandra Prasetyo Utomo

Reference:

Machine Learning course (Coursera) from Andrew Ng