15/11/2021

# Machine learning

- Supervised learning
  - Classification
    - Perceptron (binary)
  - Regression
    - Linear Regression ✓
    - logistic Regression
- unsupervised learning
  - Clustering
  - Dimensionality Reduction

## Regression (Real-valued function)

(Prediction) (Continous Data)

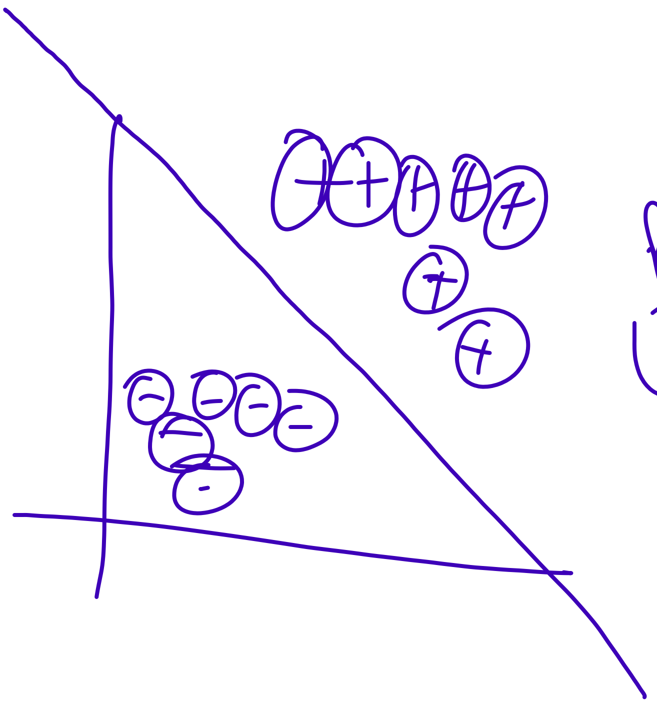
↳ House Price Prediction

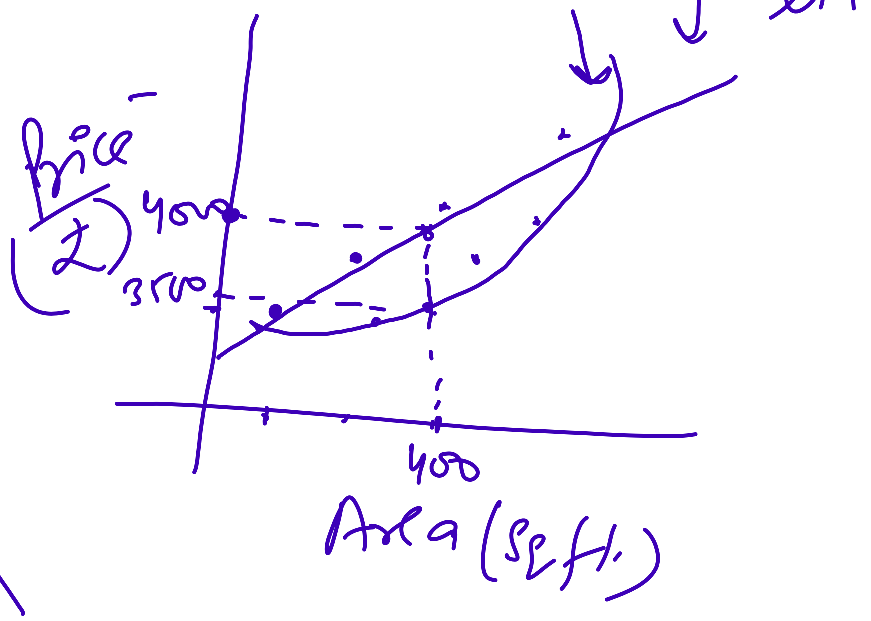
training data

| | X Area (sq. ft) | Y Price |
|---|---|---|
| 1 | 100 | 1500 ₹ |
| 2 | 200 | 2500 ₹ |
| 3 | 900 | 10000 ₹ |
| ⋮ | 600 | 6500 ₹ |
| n | 300 | 3500 ₹ |

what will be the price for 4000 sqft. (Testing data)

$$(x_i\ y_i)_{i=1}^{n}$$



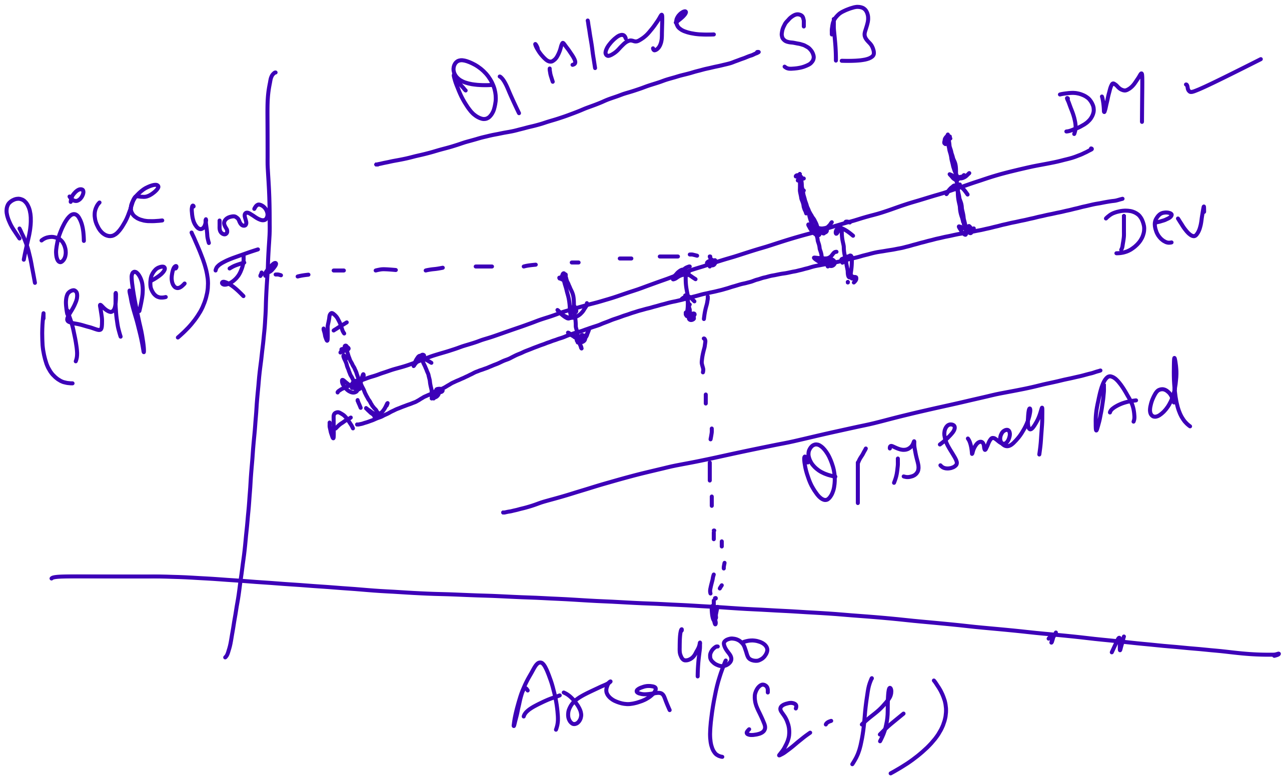



Curve → Polynomial Regression

Line → Linear Regression

## Linear Regression :-

1) Build the Model
2) Draw the line
3) Best fitted line which will try to all the training data points.




A : actual point

A' : predicted point

Prediction price → draw a line → min distance → find m & C

**Equation for line**

$y = mx + C$

Prediction price

400

**To find m & C**

m : slope of line

C : intercept

* No, this line will not give 100% guarantee to get the best prediction.

* Two functions is used for Linear Regression

① Hypothesis function $h_\theta(x)$

② Cost function $J(\theta_0, \theta_1)$

$$h_\theta(x) = mx + c$$

$\theta_1$ : intercept

$\theta_0$ : slope

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

eg $\theta_0, \theta_1$ ✓ ✓

$\theta_0 = 1$ ✓

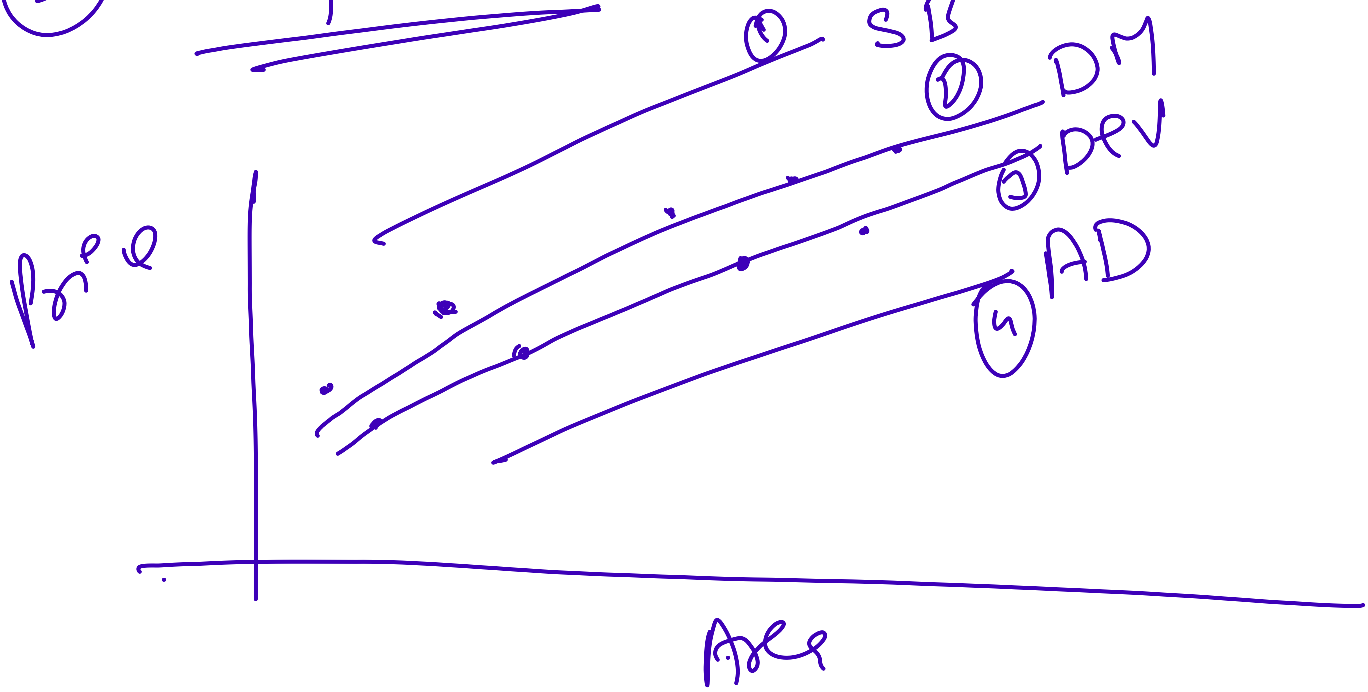
$\theta_1 = 0$ ✓

$$h_\theta(x) = 1 \times x + 0$$

$$\boxed{h_\theta(x) = x}$$

ideal case

② Cost function $(\theta_0, \theta_1)$

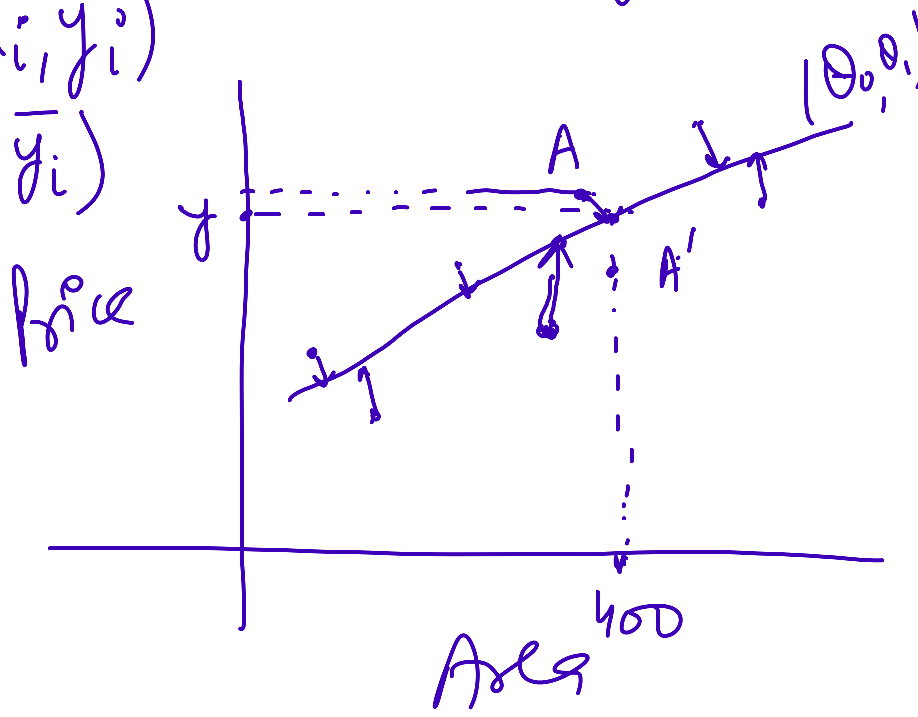



$$\underbrace{\text{Mean Squared distance}}_{\text{Cost function for linear Regression}}$$

A: Actual point $(x_i, y_i)$

A': Predicted point $(x_i, \bar{y}_i)$

$(\bar{y}_i - y_i)$ — +ve / −ve



$$\boxed{J(\theta_0, \theta_1) = \frac{1}{2n} \sum_{i=1}^{n} (\bar{y}_i - y_i)^2}$$

(MSE)

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

ex

Testing data)
4 sqft

↓

Draw line

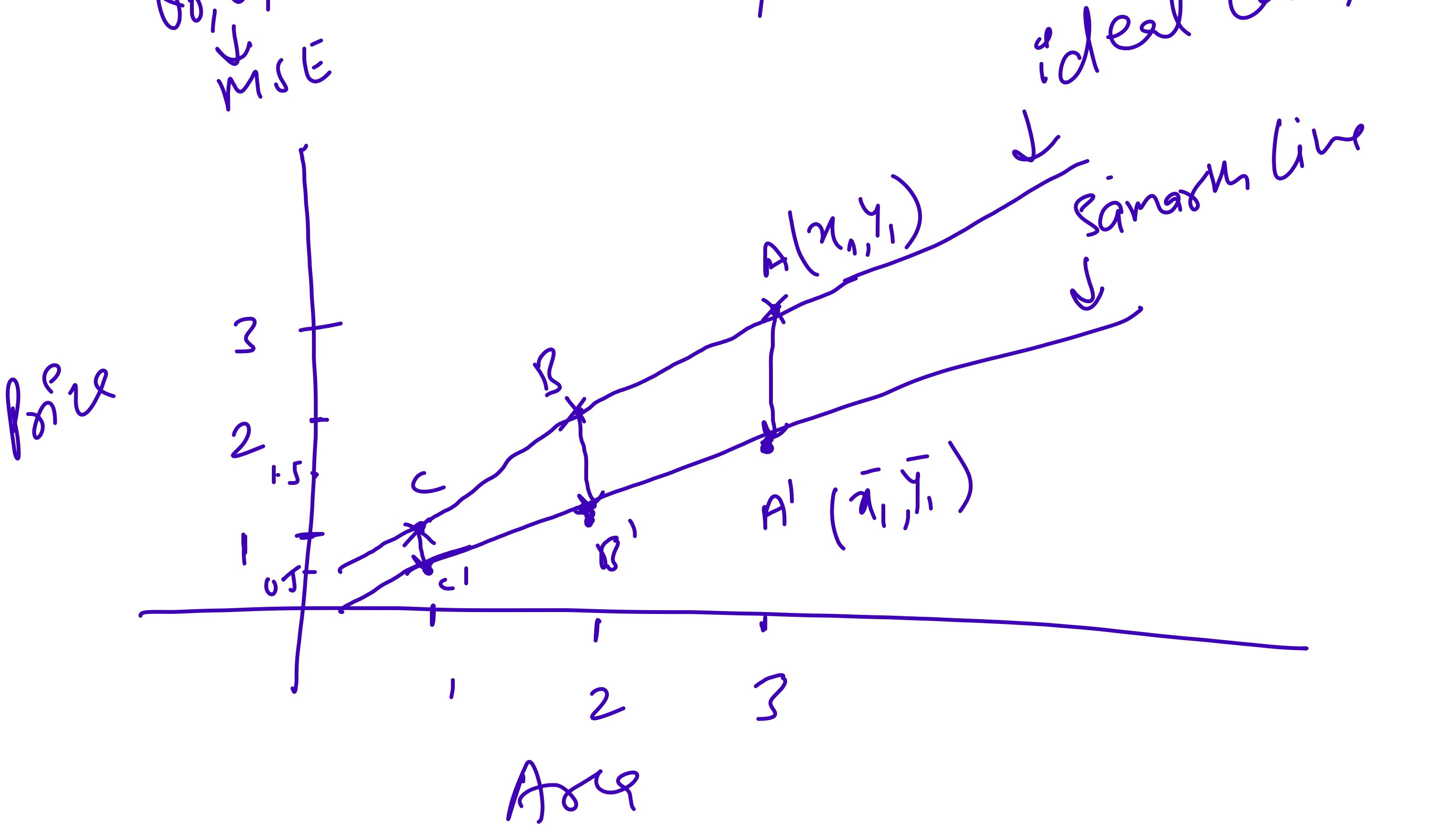
↓

$\theta_0, \theta_1$

↓

MSE

| | x Area (sq.ft) | y Price (₹) |
|---|---|---|
| $x_1$ | 1 | 1 |
| $x_2$ | 2 | 2 |
| $x_3$ | 3 | 3 |



Case 1 :-

$\theta_0 = 0.5$

$\theta_1 = 0$

$$h_\theta(x) = \theta_0 x + \theta_1$$
$$= 0.5x + 0$$
$$h_\theta(x) = 0.5x$$

$x_1 = 1$, $h_\theta(1) = 0.5$, $(1, 0.5)$

$x_2 = 2$, $h_\theta(2) = 0.5 \times 2$, $(2, 1)$

$x_3 = 3$, $h_\theta(3) = 0.5 \times 3$, $(3, 1.5)$

Cost function

$$J(\theta_0, \theta_1) = \frac{1}{2n} \sum_{i=1}^{n} (\bar{y}_i - y_i)^2$$

$n = 3$
$\theta_0 = 0.5$
$\theta_1 = 0$

$$= \frac{1}{2 \times 3} \sum_{i=1}^{3} (h_\theta(x) - y_i)^2$$

$$J(0.5, 0) = \frac{1}{6} \sum_{i=1}^{3} (h_\theta(x) - y_i)^2$$

$$J(0.5, 0) = \frac{1}{6}\left[(0.5\times 1 - 1)^2 + (1-2)^2 + (1.5-3)^2\right]$$

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

$$= 0.58$$

<u>Perided Can</u>

$$\theta_0 = 1$$

$$\theta_1 = 0$$

$$h_\theta(x) = x$$

| | | |
|---|---|---|
| $x=1$, | $h_\theta(1) = 1$ | $(1,1)$ |
| $x=2$ | $h_\theta(2) = 2$ | $(2,2)$ |
| $x=3$ | $h_\theta(3) = 3$ | $(3,3)$ |

$$J(1,0) = \frac{1}{2\times 3}\sum_{i=1}^{3}(1-1)^2 + (2-2)^2 + (3-3)^2$$

$$= 0$$

Case 2

$\theta_0 = 0$

$\theta_1 = 0.5$

G.D