25/1/25

ML

Given a dataset for performing linear regression with one variable.

Example:- Used car price prediction

| I/p → Features: Miles Driven | Engine Capacity | o/p: Price |
| --- | --- | --- |

But for simplicity purpose we are taking only one i/p.

| Miles Driven | Price |
|---|---|
| 1000 | 2.5 |
| 2520 | 3.6 |
| ⋮ | ⋮ |

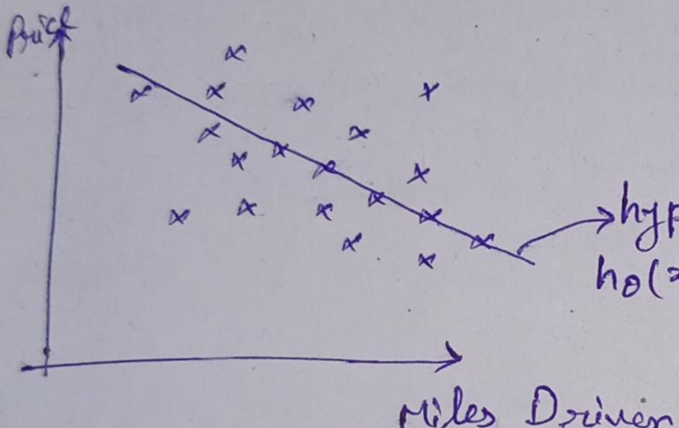
let, 2D matrix

(1000, 2.5)

(2520, 3.6)

$y = mx + c$

$= \theta_1 x + \theta_0$




→ hypothesis

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

$\Downarrow$ $\Downarrow$

Find Find

CSV (Comma Separate Vector) can store huge no. of datas in tabular form.

25/1/25

1) What is machine learning?

2) What is learning?

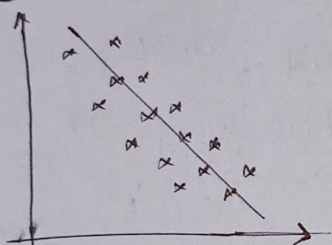
3) What are the different types of m/l? → Supervised learning, → Unsupervised learning, → Semisupervised learning

4) What is Supervised learning!

5) What type of problem we can solve using Supervised learning!
↓ Regression ↓ Classification

6) What is Regression problem explain with a real-life example?

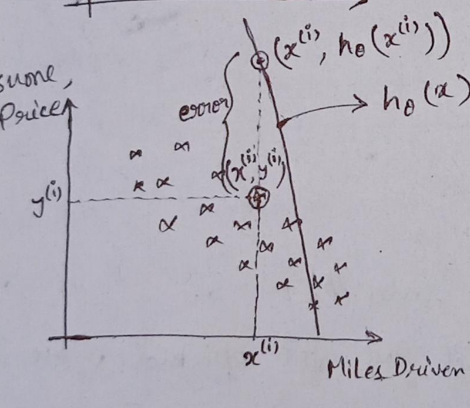
## Linear Regression with one variable



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

Initially choose $\theta_0$ and $\theta_1$ randomly.

Assume,




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

$\text{error} = \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

~~We only need +ve values~~

The distance can be -ve but we only need magnitude, but we can't do mod(| |). Because, for later we have to do derivative but mod ~~make~~ can make problem at then. So, we can do square.

Now, assume there are m data samples.

So, total no. of errors $= \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}_{\text{mean squared error (MSE)}}$ → loss function

## Minimize Error

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

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

When we have more than one variable then we need to do partial derivative.

$y = mx + c$
$\Downarrow$
slope

$$m = \frac{dy}{dx}$$

$$\frac{dJ(\theta_0, \theta_1)}{d\theta_0} \qquad \frac{dJ(\theta_0, \theta_2)}{d\theta_1}$$

Updating the value of $\theta_0, \theta_1$ :

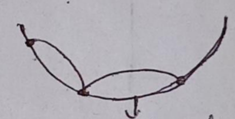
$$\theta_0 = \theta_0 - \alpha \frac{dJ(\theta_0, \theta_1)}{d\theta_0} \rightarrow \text{gradiant}$$

$$\theta_1 = \theta_1 - \alpha . \frac{dJ(\theta_0, \theta_1)}{d\theta_1}$$

This is called Gradiant Decent Algorithm for learning.

$\alpha$ = Learning rate

If $\alpha = 1$ then it means $\alpha$ is too high. That means it is seeing the slope and making a high jump. It can make problem.




Steps :

(i) Take Random points
(ii) Calculate errors
(iii) Calculate gradiants
(iv) Calculate $\alpha$ - value
(v) Calculate new value for $\theta_0$ and $\theta_1$
(vi) Repeat from step (i) until we get optimal value

Performing Partial Derivative:

For $\theta_0$

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

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

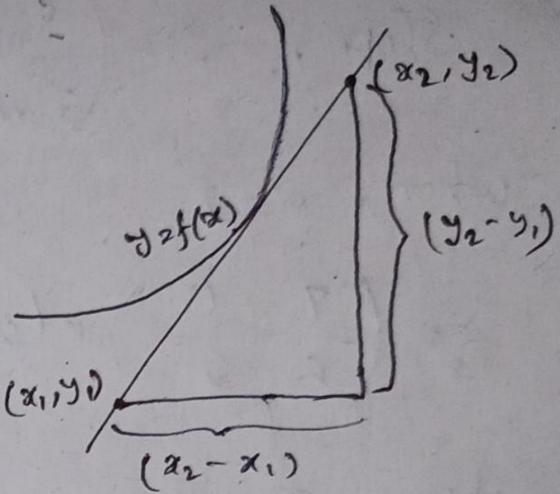
$$\frac{d}{d\theta_0} J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \frac{d}{d\theta_0} \left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)^2$$

$$= \frac{1}{2m} \sum_{i=1}^{m} \frac{d\left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)^2}{d\left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)} \times \frac{d\left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)}{d\theta_0}$$

$$= \frac{1}{2m} \sum_{i=1}^{m} 2\left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)(1 + 0 - 0)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)$$

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




Now, for $\theta_1$

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

$$(x + dx)^2$$
$$= x^2 + x\,dx + x\,dx + dx^2$$
$$\Rightarrow (x + dx)^2 - x^2 = 2x\,dx$$
$$\Rightarrow \frac{(x + dx)^2 - x^2}{dx} = 2x$$

$$\tan\theta = \frac{\text{height}}{\text{base}}$$
$$= \frac{y_2 - y_1}{x_2 - x_1}$$
$$= \frac{dy}{dx}$$