Machine Learning

MTE - 20 } Practical / Implementation
ETE - 40 }

PRS - 25

CWS - 15

→ Attendance : 5

>95	5
90-94	4
85-89	3
80-84	2
75-79	1
<75	0

→ Test : 2

Assignment : 1

Pre-requisite:

Probability

Differentiation

$$\frac{\partial}{\partial x}(2) = 0$$

$$\frac{\partial (x^n)}{\partial x} = n x^{n-1}$$

GG
GB
BB
BG

$\frac{1}{3}$

ML ?

Data → Patterns → External data

$\underbrace{\text{Data} \rightarrow \text{Patterns}}_{\text{Data Analytics}}$

Training Data

D →
H →
H →
H →
D →

ht > 4ft → H
ht < 4ft → D

New Image → Testing Data

Why?

- Data
- Computations
- Algorithms/Mathematical Model

ML

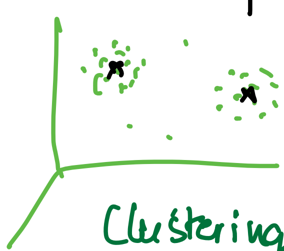
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Supervised Learning

Input → label

→ Cat
→ Horse
→ Dog

Unsupervised Learning



Clustering

Reinforcement Learning

Semi-supervised

ht color legs eye color → features x

Cat → label y^1 → x^1
Dog → y^2 → x^2
Horse → y^3 → x^3
⋮

m data points

$x_1^1 \quad x_2^1 \quad x_3^1 \quad x_4^1 \quad \dots \quad x^m$

Supervised Learning : $\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

$x^{(i)} = \{x_1^i, x_2^i, \dots, x_n^i\}$ → x^i has n features

$x^i \in R^n$

$\boxed{x_j^i = \text{jth feature of ith example}}$

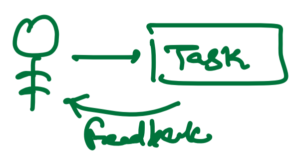
Features

- Numerical (continuous)
- Categorical {Black, Brown, Green, Grey}

Label

- Numerical (Continuous) — house price prediction (Regression)
- Categorical {Cat, Dog, Horse} (Classification)

Reinforcement



Reward
Punish

Environment Situation

LINEAR REGRESSION

Eg:

Training Data

Time Spent (Feature)	Marks (label)
1	4
3	7
10	8
20	10

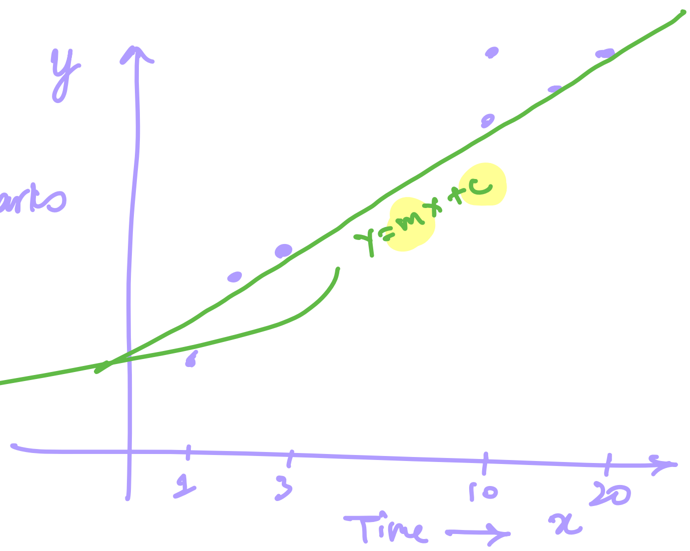
$x^{(i)} \in R$

$y^{(i)} \in R$

→ MODEL ? (Pattern)

Q: 8hrs ? Score ? } Test Data

?



$Y = mx + c$

$= m \cdot 8 + c$

Hypothesis $Y = \theta_1 x + \theta_0$

x_1	x_2	x_3	x_4
Hours	CGPA	Classes	Assign

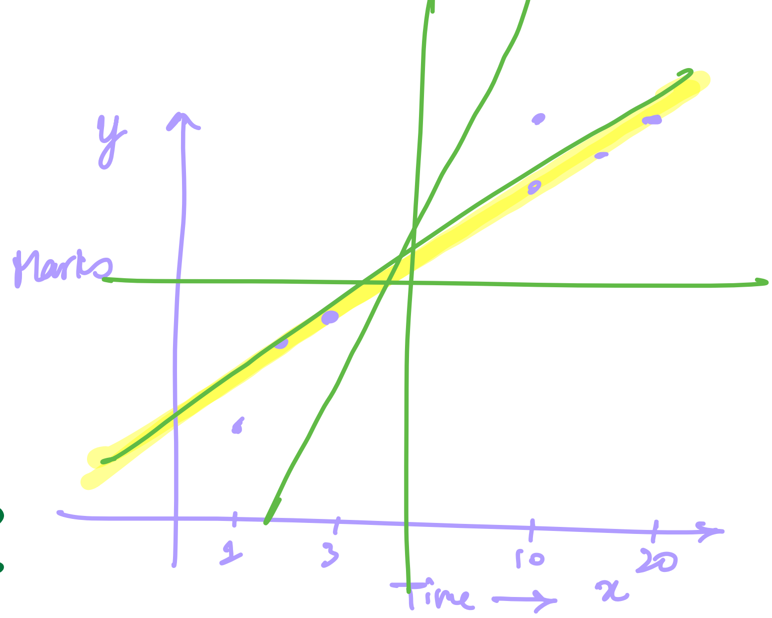
$Y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_0$

$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$

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

Aim:

To learn best line which fits through data points

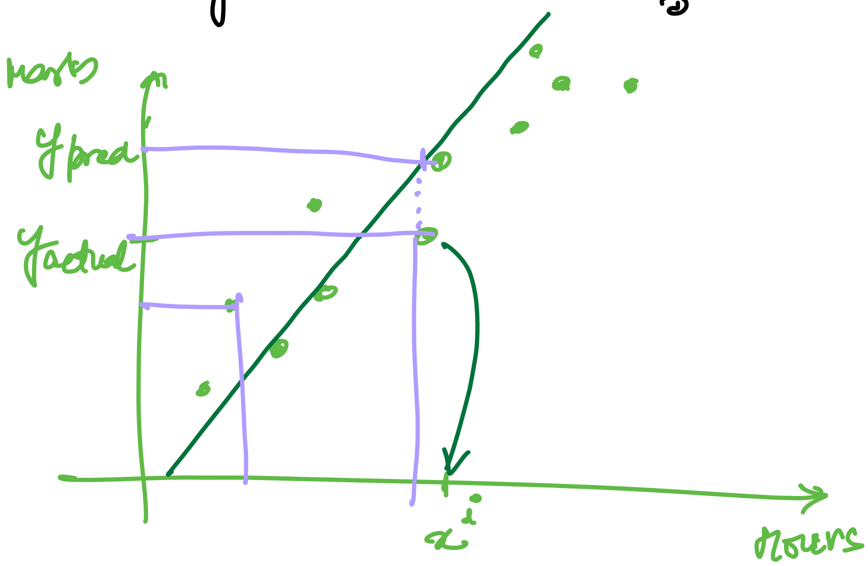


(a) Random value of Θ start

(b) How good the line is ?

(c) Θ change/update good performance

How good our Θ is ?



$$\epsilon^{(i)} = \left| y^{(i)}_{pred} - y^{(i)}_{actual} \right|$$

Error for i^{th} example

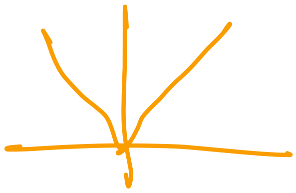
$$\text{Total Error for all points} = \sum_{i=1}^{m} \left| y^{(i)}_{pred} - y^{(i)}_{actual} \right|$$

$$\text{Total Error for all points} = \sum_{i=1}^{m} \left| \hat{y}^{(i)} - y^{(i)} \right|$$

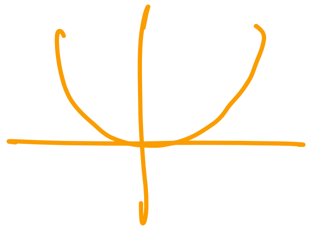
$$\text{Average Error} = \frac{1}{m} \sum_{i=1}^{m} \left| \hat{y}^{(i)} - y^{(i)} \right|$$

(Average Absolute Error)

$|x|$



x^2



Total Error for all points $= \sum_{i=1}^{m} \left[\hat{y}^{(i)} - y^{(i)}\right]^2$

Mean Squared Error (MSE) $= \frac{1}{m} \sum_{i=1}^{m} \left[\hat{y}^{(i)} - y^{(i)}\right]^2$

predicted

actual

Loss function

J

minimize

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

- make updation to your θ, so that it becomes a better θ

$y = (x-5)^2$

For what value of x y will be minimum?

$$\frac{dy}{dx} = \frac{\partial (x-5)^2}{\partial x} = 2(x-5) = 0$$

$x = 5$

Gradient Descent