DATE __/__/__ ML-1

Introduction to machine learning.

Agenda -

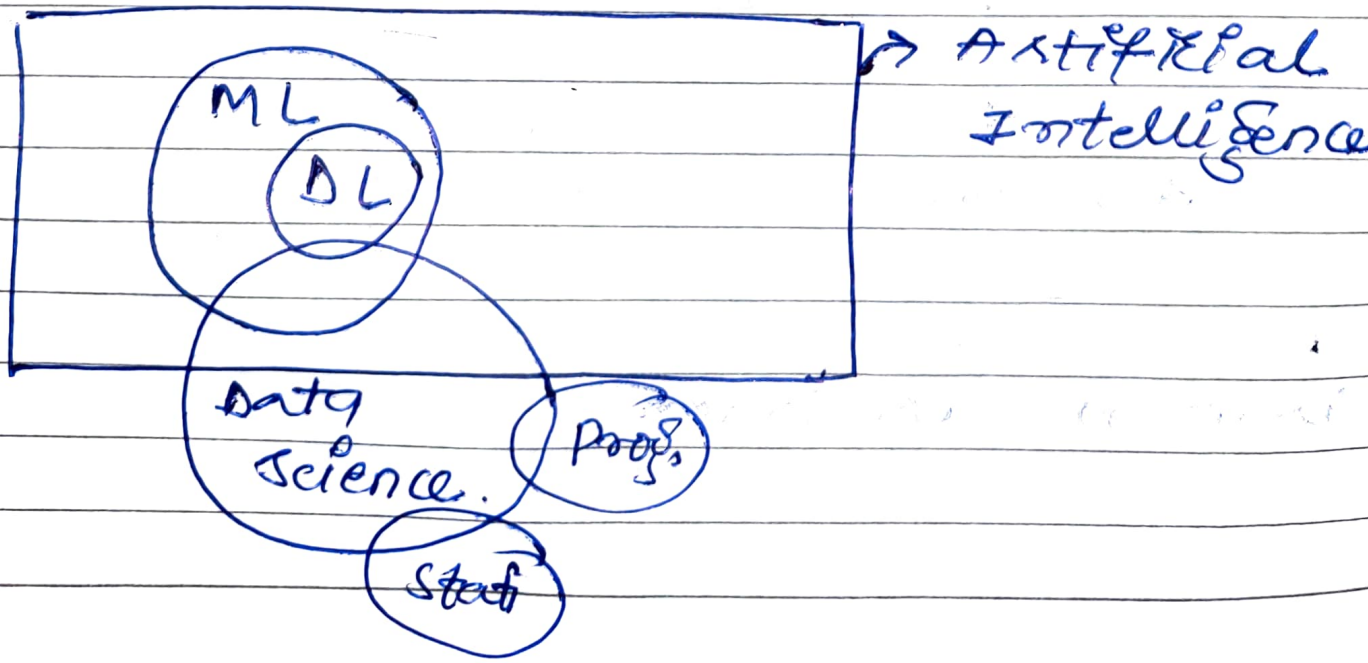
1) ML Introduction
2) AI VS ML VS DL VS DS
3) Simple ~~Learning~~ linear regression → mathematical Int.

1) AI VS ML VS DL VS DS.

streaming platform

AI → [netflix] ← Recommendation system.

Application → AI module



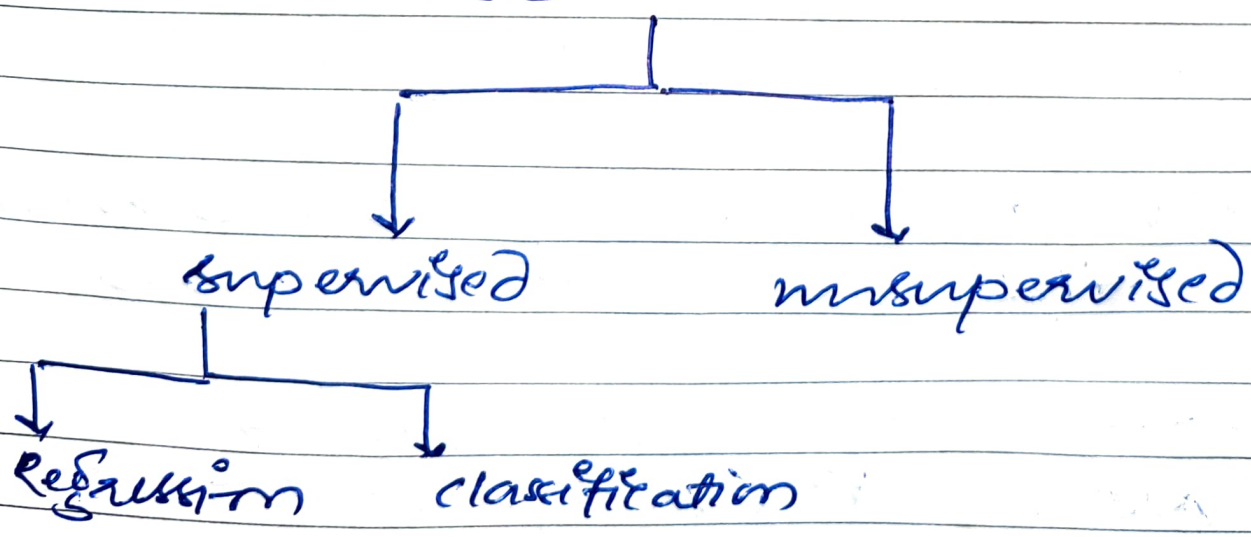
DATE ___/___/___

AI – Artificial intelligence it is creating an application where it performs all its task without any human interaction.
ex – Alexa, chatbot, self driving car.

ML – It provides stat tools to explore analyse, visualize and perform prediction and ~~other~~ task with the help of data.

DL – uses multilayer neural network which mimics the human brain.

ML and DL



DATE ___/___/_____

Regression -

1) linear Regression
2) logistic Regression
3) polynomial
4) SVM → classification
5) SVR
6) Decision tree
7) Random forest
8) xg-boost
9) KNN

unsupervised

clustering algorithm

DBmeans Kmeans Hearchical

supervised learning → we know the output variable

unsupervised learning → we don't know about our output variable

DATE ___/___/___

supervised

ex 1 –

Degree	Exp	salary
B.E	7	50k.
PHD	2	70k.
–	–	–
–	–	–
=	=	=

↓

→ output variable

→ if it is continuous so we use regression.

ex 2 –

no. of play hrs	no. of study hours	Pass/fail
9	1	0
7	2	0
3	5	1

↓ dependent feature

Independent feature

Here we have two categories 0, 1 so this is categorical feature.

DATE ___/___/_____

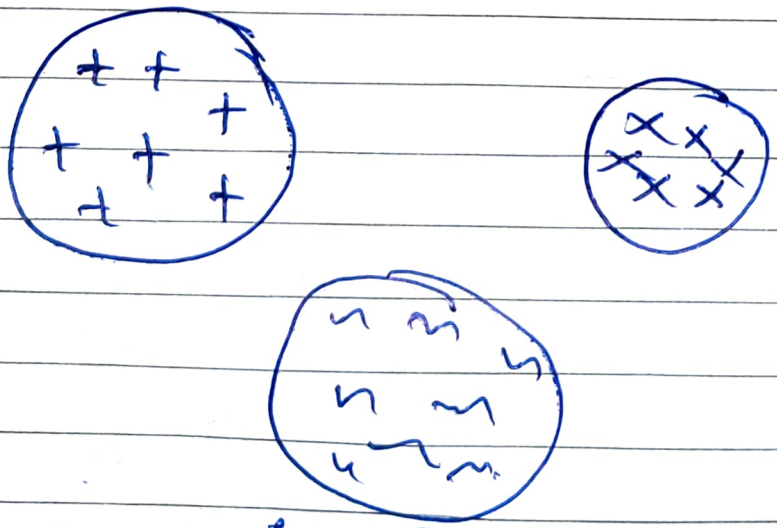
In simple linear regression -

Age	salary	spending-score (1-10)
24	70k	1
26	100k	9
-	-	-
21	20k	9
25	120k	2

product

10%

20%



clustering our customer.
customer segmentation.

DATE / /

1) Simple Linear Regression.

1 - Independent feature
1 - dependent feature

DATABASE

Height	weight

Aim - To create a model
I/P → Height O/P → weight

Ex-2 DATASET

X	Y
no. of rooms	price

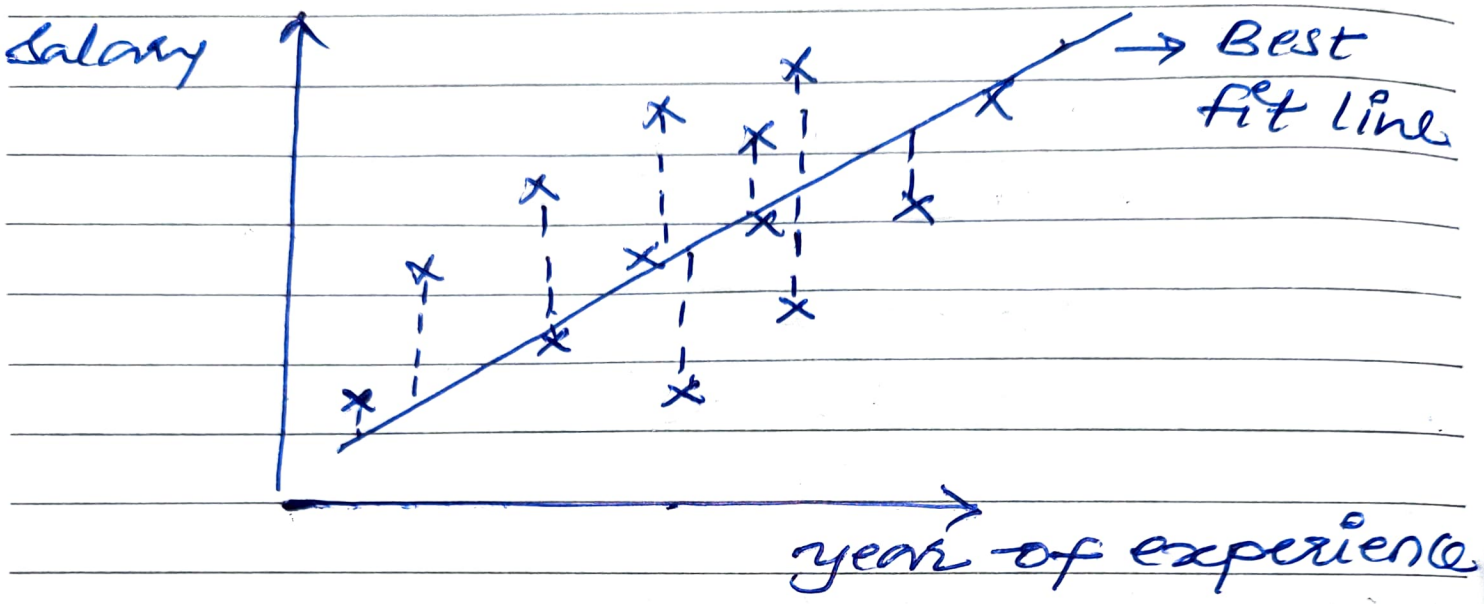
Ex-3

X	↑
year of Exp	salary

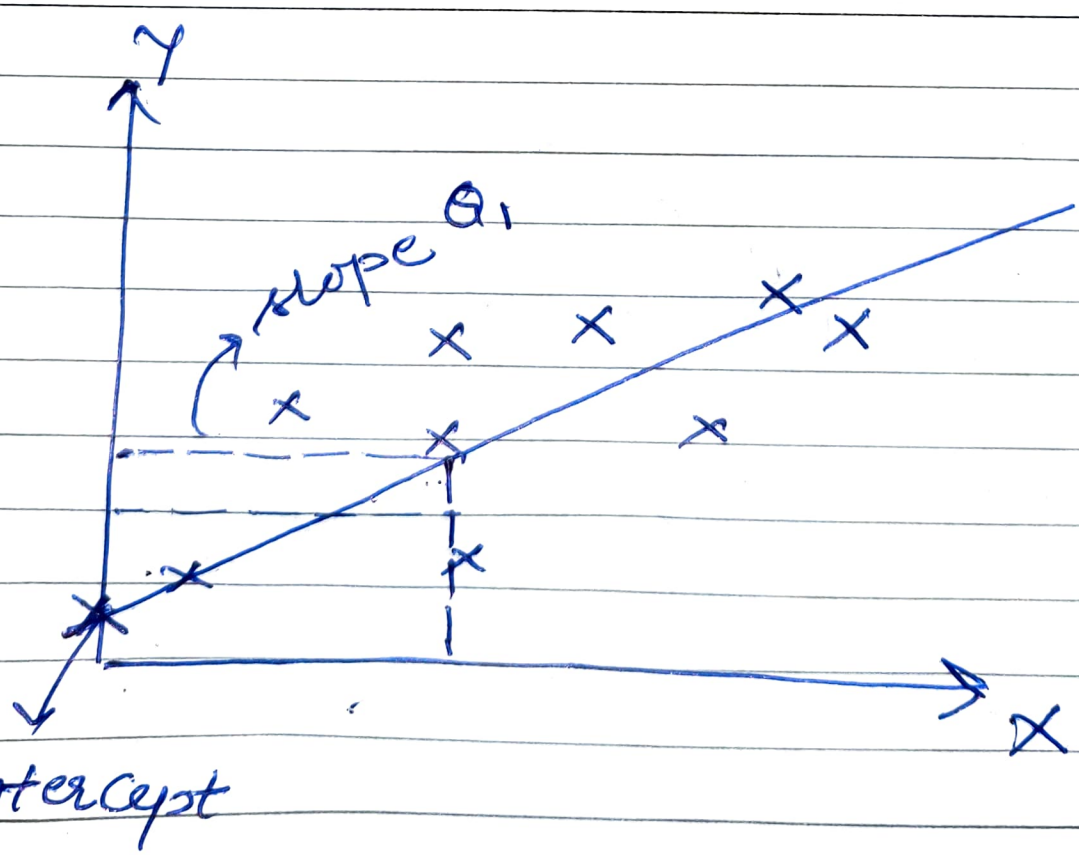
supervised and Regression problem.

model → year of experience & salary
predict → salary based on I/P year

DATE ___/___/_____



Best fit line - the error is minimum, the error residual between original point and predicted point should be minimum.



DATE ___/___/___

equation of straight line

$$y = mx + c$$
$$y = \beta_0 + \beta_1 x_1$$

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

θ_0 ↓ intercept, θ_1 ↓ slope

to get the best fit line we tweak the value of θ_0 & θ_1

to minimize the residual error we use cost function.

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

$h_\theta(x)^{(i)}$ ↓ predicted, $y^{(i)}$ ↓ Actual

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

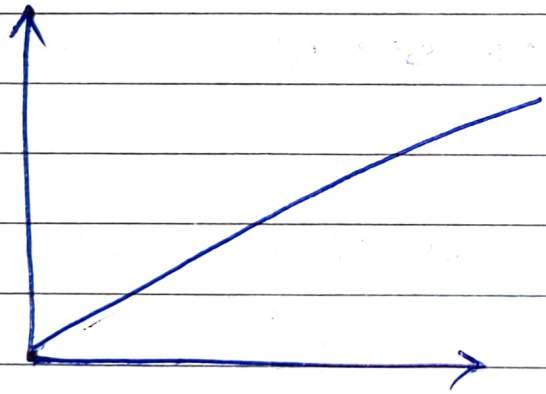
↳ mean square error.

DATE ___/___/_____

Final Aim -

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

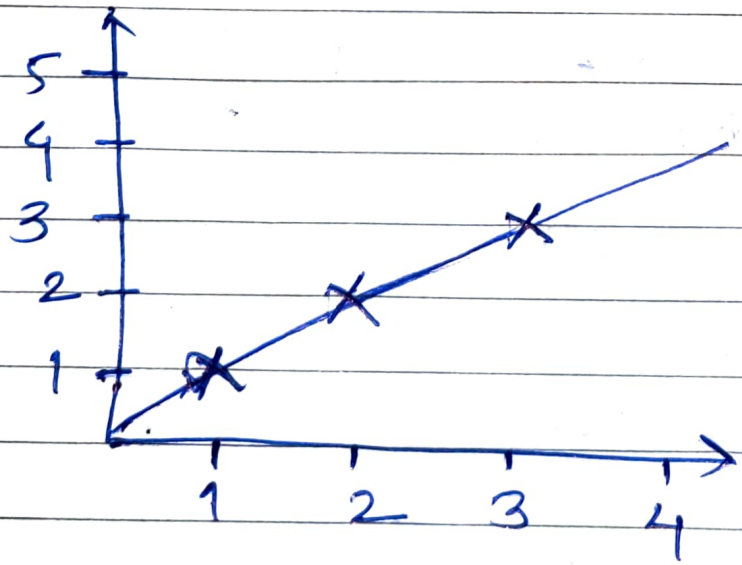
$$h_\theta(x)^{(i)} = \theta_0 + \theta_1 x$$



let's consider.

$\theta_0 = 0$

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



$h_\theta(x) = \theta_1 x$.

$\theta_1 = 1$

$h_\theta(x) = 1 \quad x = 1$

$h_\theta(x) = 2 \quad x = 2$

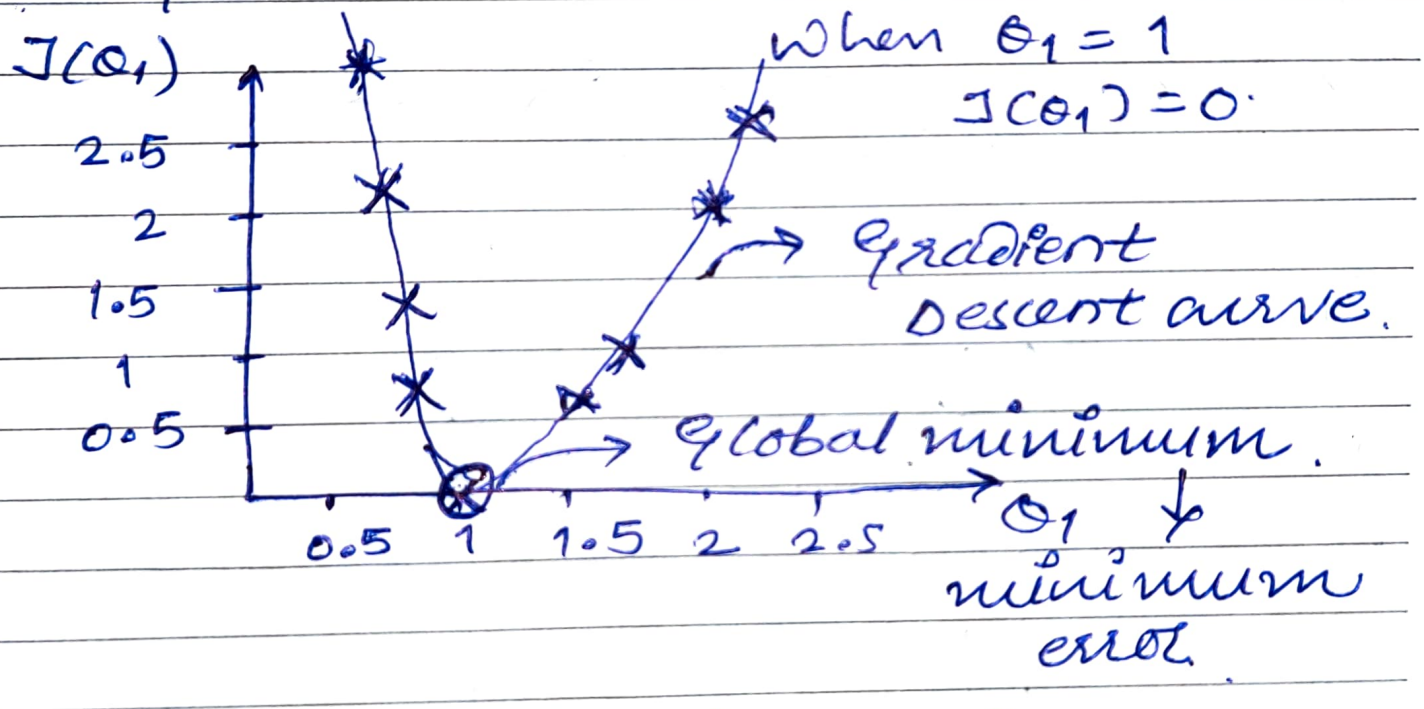
$h_\theta(x) = 3 \quad x = 3.$

1.5
DATE 7.5

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

$$= \frac{1}{3} [0 + 0 + 0] = 0$$

example



$\theta_1 = 0.5$ $\quad h_\theta(x) = \theta_1 x$.

$h_{\theta_1}(x) = 0.5 \quad x = 1$

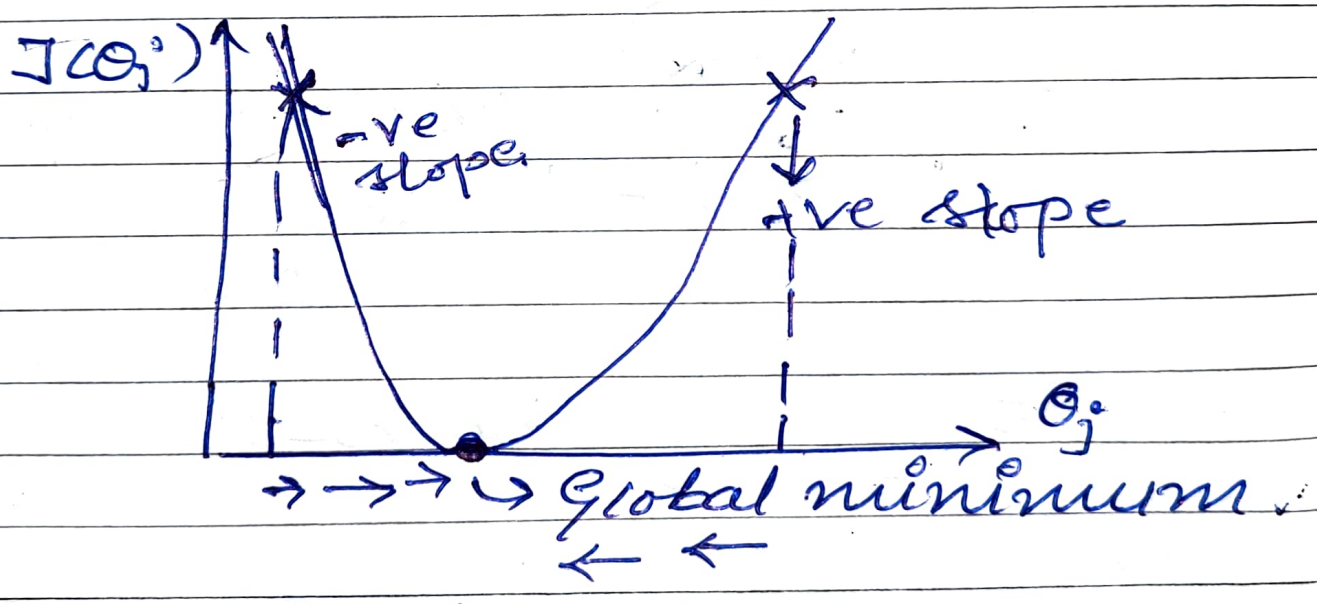
$h_{\theta_1}(x) = 1 \quad x = 2.$

$h_{\theta_1}(x) = 1.5 \quad x = 3$

DATE ___/___/___

convergence Algorithm -
optimize the changes of θ_j value.

Repeat until convergence
{
$\theta_j^\circ = \theta_j^\circ - \alpha \frac{\partial J(\theta_j)}{\partial \theta_j}$ → slope.
↳ learning rate
}



$\theta_j^\circ = \theta_j^\circ - \alpha(-ve)$.
$= \theta_j^\circ + \alpha$

$\theta_j^\circ = \theta_j^\circ - \alpha(+ve)$
$= \theta_j^\circ + \alpha$

DATE ___/___/___

learning rate-

How fast the convergence Algorithm will work.

- If α is too low then it will take too much time to reach the global minimum.

- If α is too big then it will escape the global minimum.