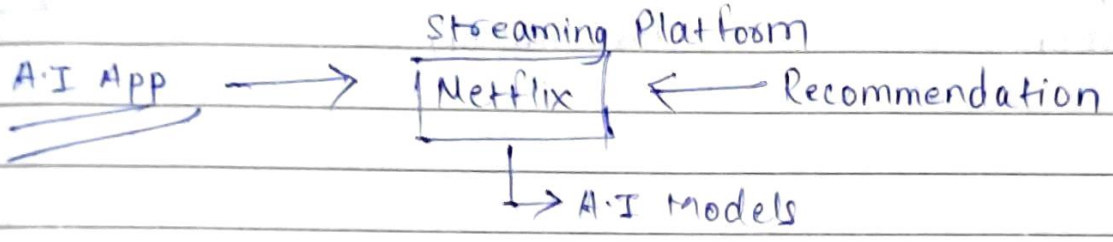
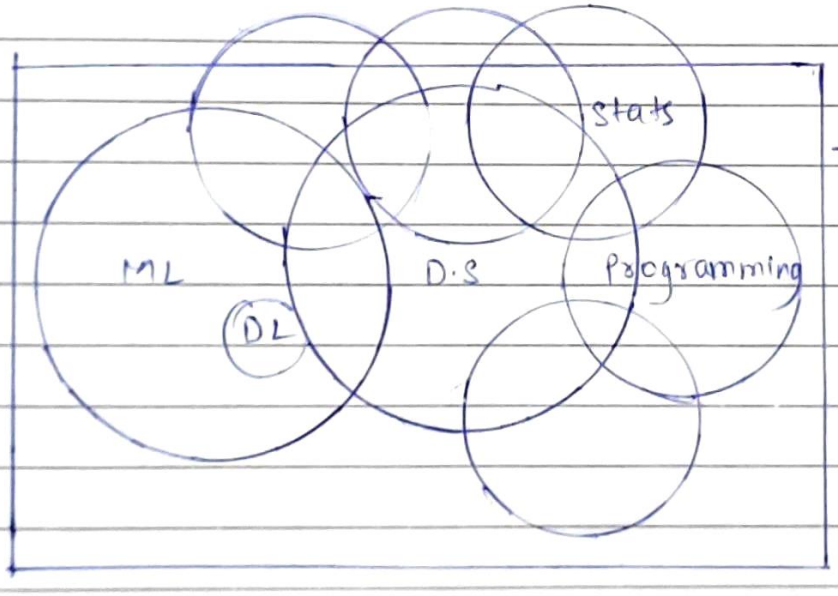
Date.....

Expt. No..... Page. No.....

Introduction to Machine learning

1. Machine learning Introduction
2. AI VS ML VS DL VS DS
3. Simple linear Regression → Mathematical

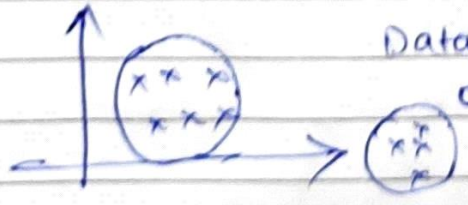




→ Artificial Intelligence (A.I)

It is an Application where it performs all its task without any human intervention

ML Provides stats to analyze / visualize / perform prediction and other task with the help of data

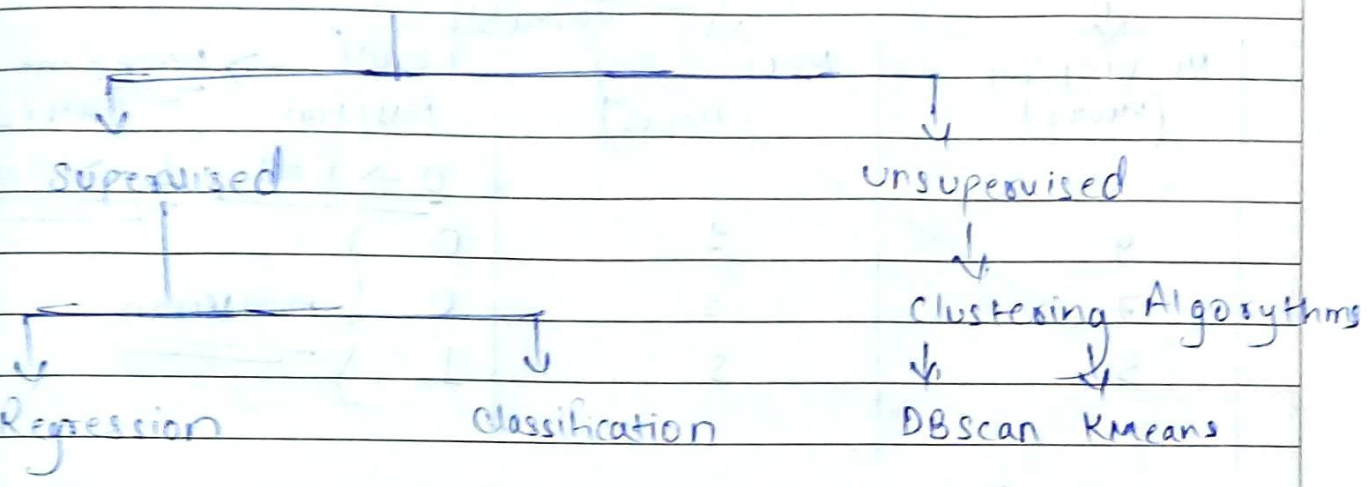


Teacher's Signature.....

Date.....

Expt. No.....

Page. No.....



1. Linear Regression
2. Polynomial
3. SVR
4. Decision Tree
5. KNN
6. Xgboost
7. Random forest

1. Logistic Regression
2. SVM
3. Decision Tree
4. Random forest
5. KNN

Supervised

Ex:-

Degree	Exp	Predict Salary	O/P
B.E	7	50K	
PHD	2	70K	
-	-	65K	Regression.
-	-	66K	
-	-	68K	

Teacher's Signature.....

Examples:- Independent Variable, Prediction

No. of plays hours	Study hours	Result Pass/fail (O/P)
9	1	0
7	2	0
3	5	1

Pass → 1
fail → 0

→ Dependent feature

Dataset Examples :-

Flight Price Prediction → Regression

Algerian fire forest ⟷ classification

Air Quality Index → Regression.

Unsupervised ML →

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

* By this Dataset we can Assume that person with higher salary has higher spending score, or person with lower salary has less spending score, there may some person with lower salary has higher spending score.

Dfn: Unsupervised learning, also known as unsupervised M.L uses machine learning algorithms to analyze and cluster unlabeled datasets.

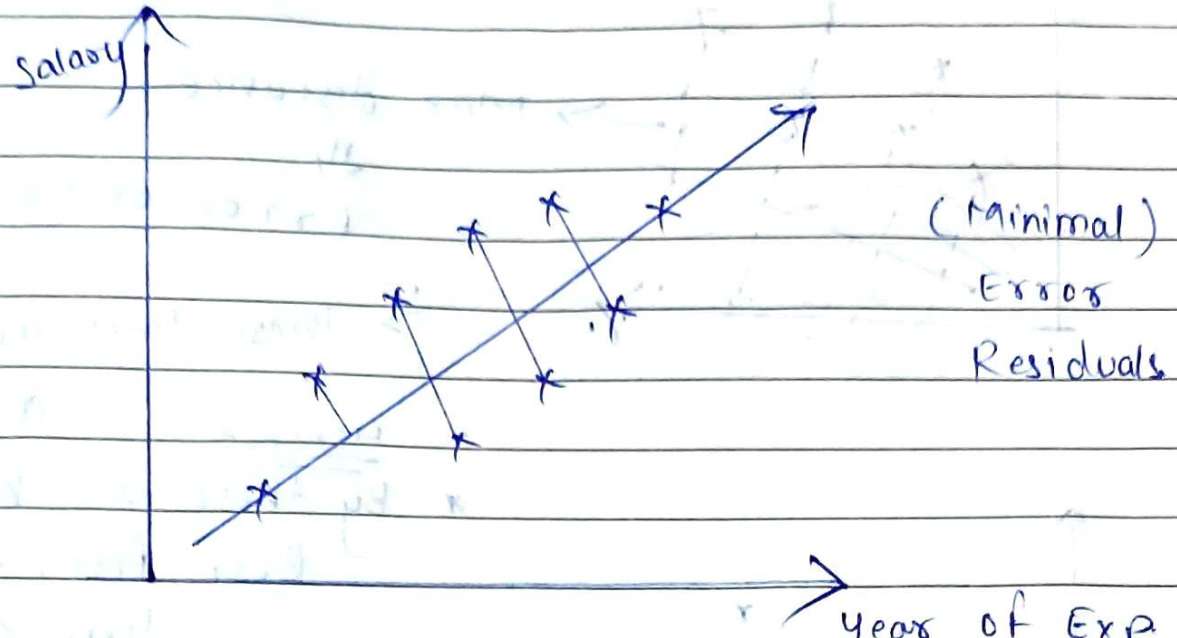
Teacher's Signature.....

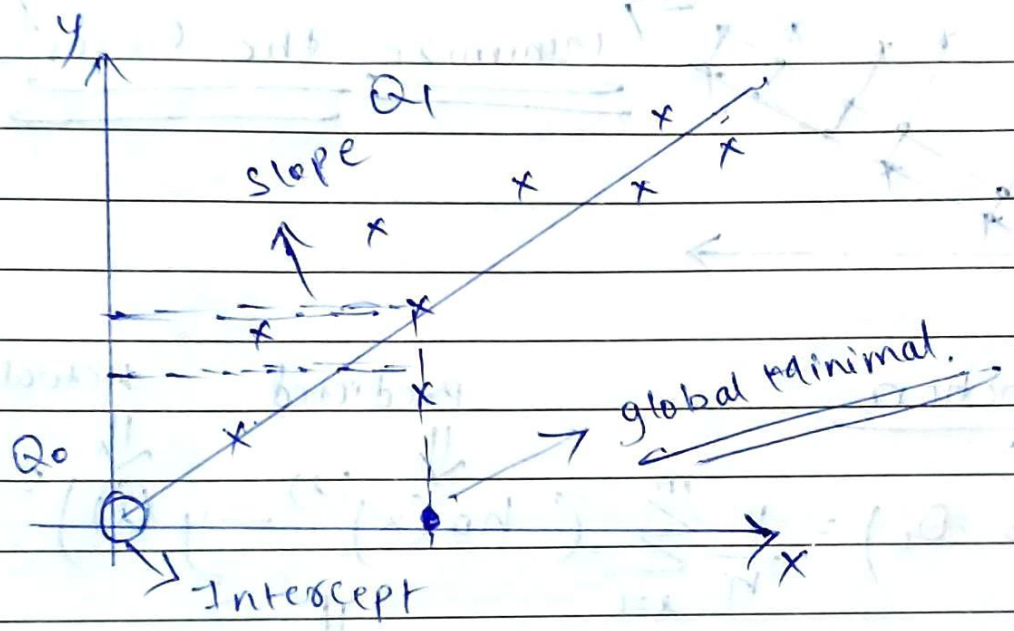
Simple Linear Regression :

1 Independent variable  1 Dependent variable 

Dataset :

Year of Exp (Independent)	Salary (O/P prediction, Dependent)
—	—
—	—
—	—
—	—
—	—





Equation of a straight line

$$y = mx + c$$

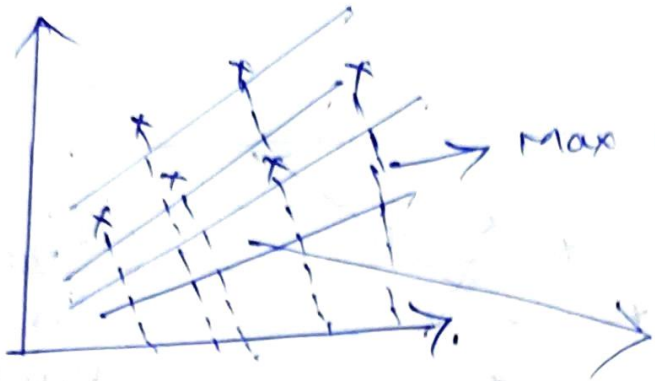
$$y = \beta_0 + \beta_1 x_1$$

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

θ_0: intercept

θ_1: slope

Teacher's Signature.....



Max distance

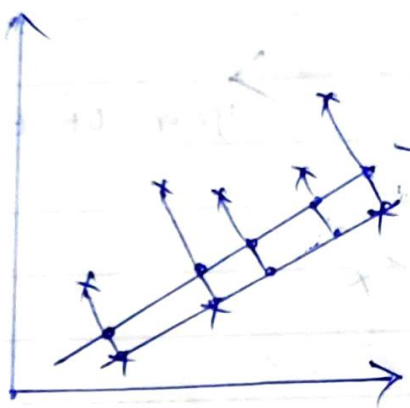
⇓

Error or Residual

These lines are training a model

Hence :

* By these we have find Best fitting straight line



Minimize the Errors

Cost Function

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

($h_\theta(x)^{(i)}$: Predicted ⇓ ; $y^{(i)}$: Actual Value ⇓)

⇓

(Mean squared Errors)

MSE

* MSE final Aim :

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

(θ_0 : intercept ↑ ; θ_1 : slope →)

formula MSE $= \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

n = number of data points

y_i = observed values

$\hat{y}_i$ = predicted values.

Examples :- $h_\theta(x) = \theta_0 + \theta_1 x$

let us consider

$\theta_0 = 0$

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

$$= \frac{1}{3}\left[(1-1)^2 + (2-2)^2 + (3-3)^2\right]$$

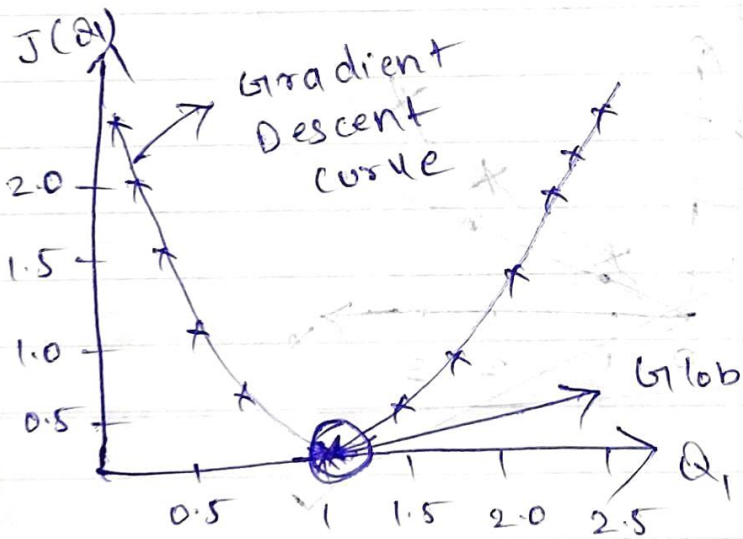
$$= 0$$

Let consider, $\theta_1 = 0.5$

$h_\theta(x) = 0.5 \quad x = 1$

$h_\theta(x) = 1 \quad x = 2$

$h_\theta(x) = 1.5 \quad x = 3$



$$J(\theta_1) = \frac{1}{3}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right.$$

$$= \frac{1}{3}\left[0.25 + 1 + 2.25\right]$$

$$= \frac{3.5}{3} = 1.16$$

Date.....

Expt. No.....

Page. No.....

$$J(\theta_1) = \frac{1}{3}\left[(0-1)^2 + (0-2)^2 + (0-3)^2\right]$$

$$J(\theta_1) \approx 4.66.$$

Convergence Algorithm { optimize the changes of θ_1 value }

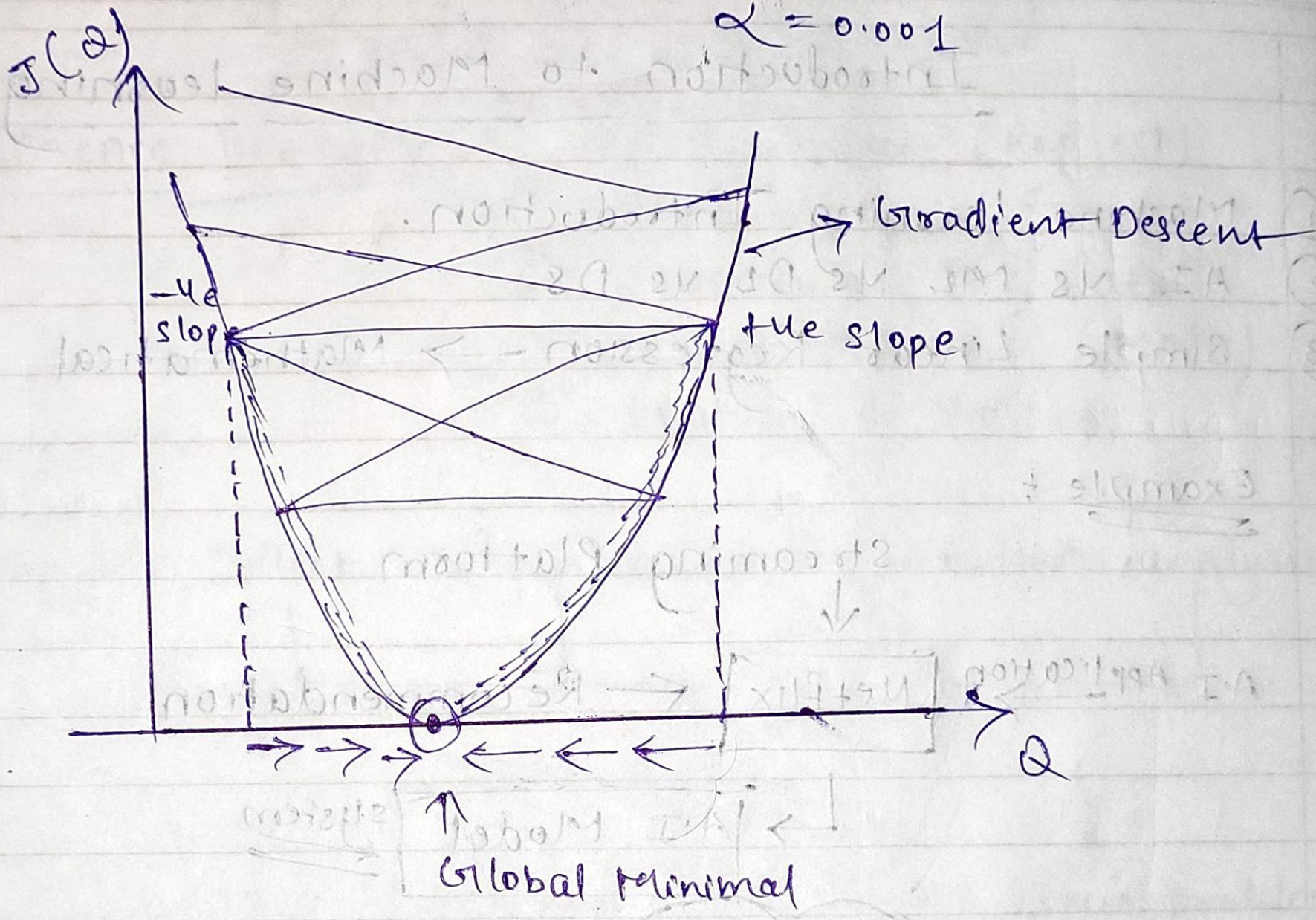
Repeat until convergence

{

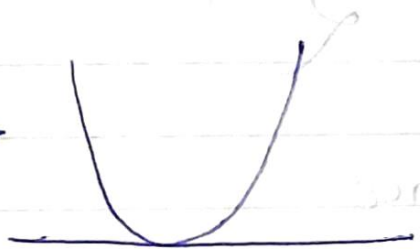
$$\theta_j := \theta_j - \alpha \left[\frac{2}{2\theta_j} J(\theta_j)\right]$$

(α → learning Rate; $\frac{2}{2\theta_j} J(\theta_j)$ ⇒ slope (Derivative))

}

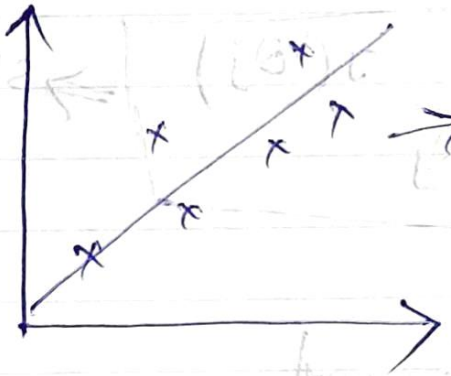

$\alpha = 0.001$
J(Q)
Gradient Descent
-ve slope
+ve slope
Q
Global minimal

MSE ⇒



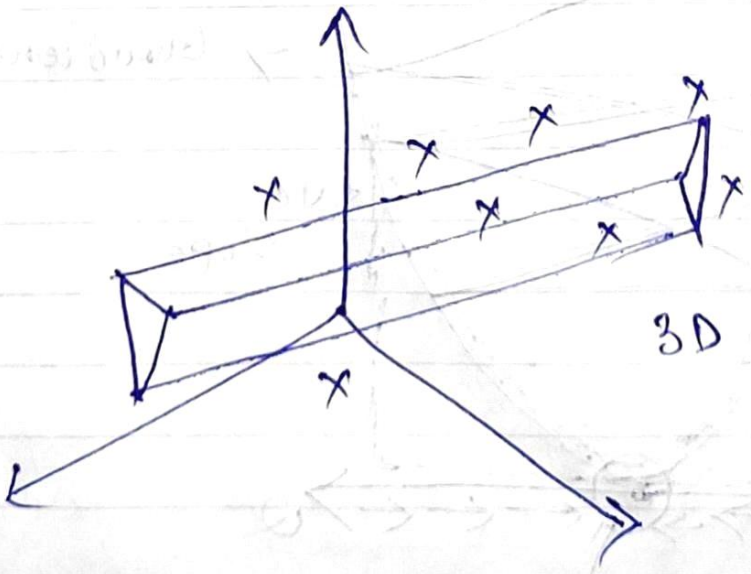
MAE

RMSE



→ Best. line fit

2D



hyperplane

3D