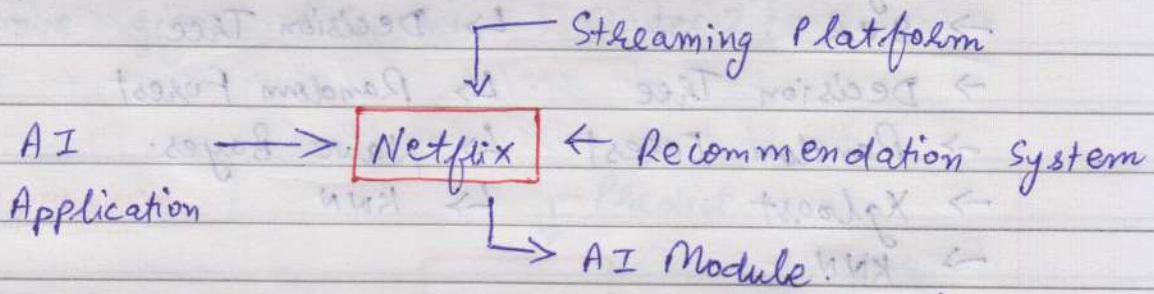
Introduction To Machine learning:- 08/10/2022

Agenda:-

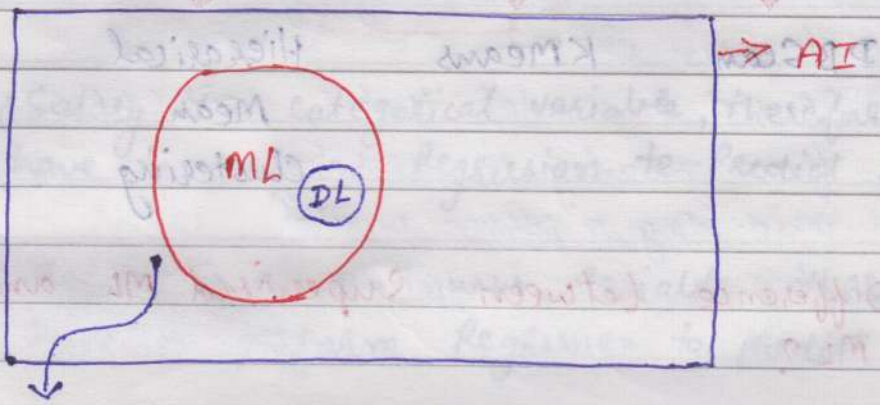
1. Machine Learning Introduction.
2. AI VS ML VS DL VS DS
3. Simple Linear Regression → Mathematical Intuition

⇒ AI Vs ML vs DL vs DS



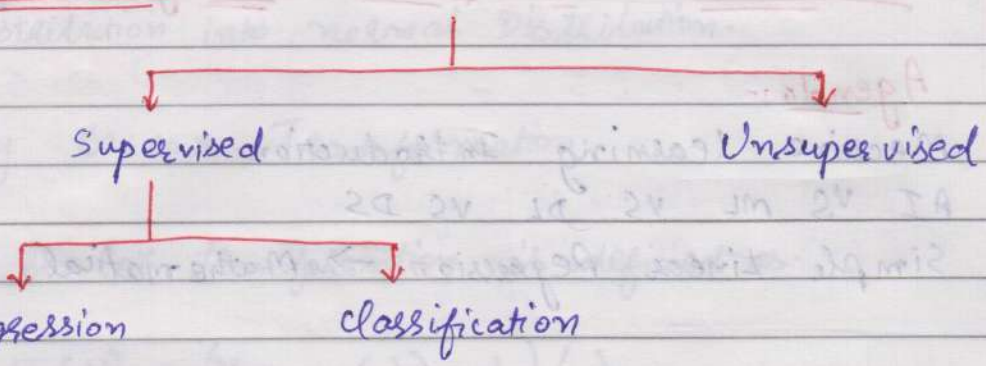
Artificial Intelligence:-

It is creating an application where it performs all its task without any human intervention.



Machine Learning provides stats tools to analyse, visualize, perform prediction and other task with the help of data.

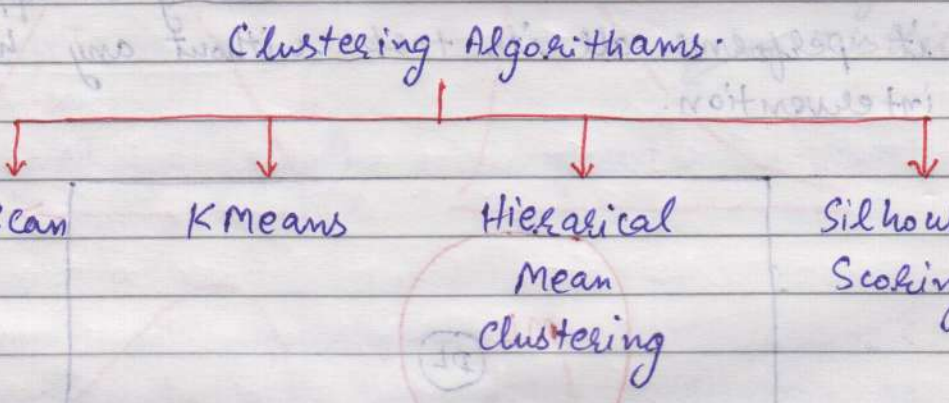
ML And DL



- → Linear Regression
- → Polynomial
- → SVR
- → Decision Tree
- → Random Forest
- → Xgboost
- → KNN

- L> Logistic Regression.
- L> SVM
- L> Decision Tree
- L> Random Forest
- L> Naive Bayes.
- L> KNN

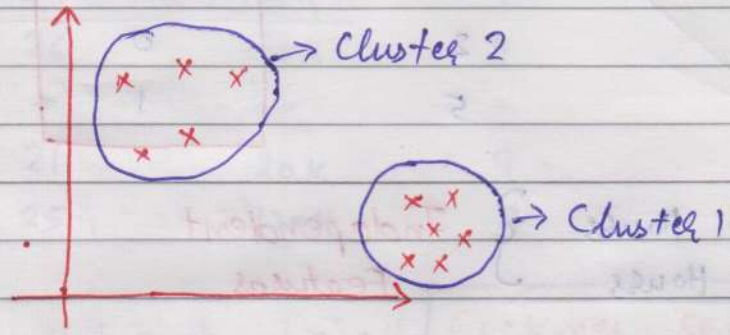
Unsupervised ML



Difference between Supervised ML and Unsupervised ML ?

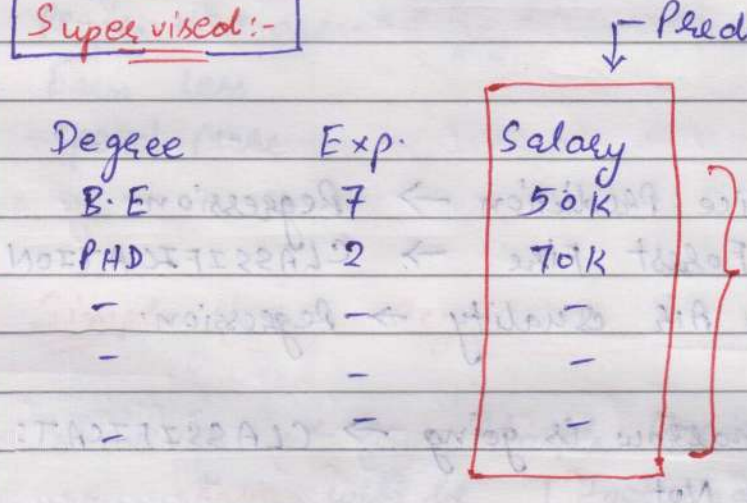
- In Supervised ML, suppose we have the dataset, we know the output variable. [Target variable]
- In Unsupervised ML, we will never know the output variable, therefore we cluster similar kind of

datapoints



In Unsupervised Learning, we do not find the output variable, therefore we make clusters like above figure [Cluster 1, Cluster 2]

Supervised:-



Note:- Salary is a continuous variable, therefore we have to perform Regression to predict Salary.

Playing Hours	Study Hours	Pass/Fail
9	1	0
7	2	0
3	5	1

classification

Here Playing Hours, Study Hours } Independent Features

Pass/Fail } Dependent Feature.

Note:- Here Pass/Fail is a categorical variable, therefore we have to perform classification, to predict the student will Pass/Fail.

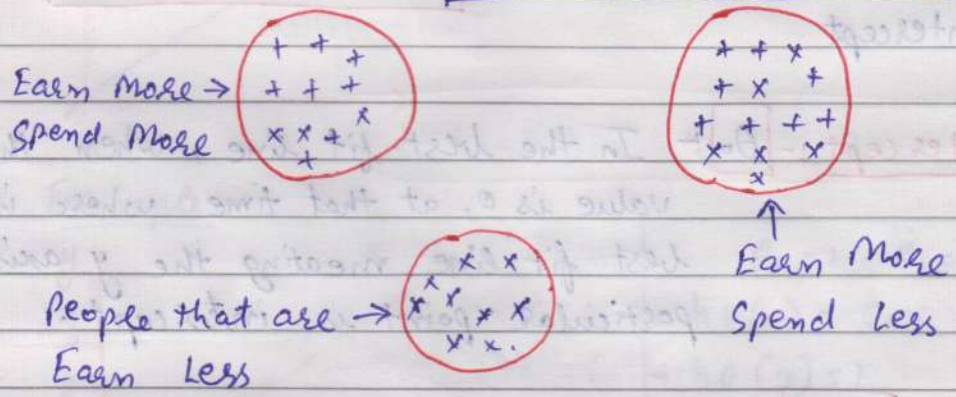
Examples:-

① Flight Price Prediction → Regression.
② Algerian Forest Fire → CLASSIFICATION
③ Predict the Air Quality Index → Regression.
④ Predict Tomorrow is going to rain or Not → CLASSIFICATION
⑤ Predict a model that from monday to saturday which day a person would buy → CLASSIFICATION

Unsupervised ML.

Age	Salary	Spending Score (1-10)
24	70K	1
26	100K	9
-	-	-
21	20K	9
25	120K	2

Customer Segmentation.



Simple Linear Regression:-

Whenever we say Simple Linear Regression, that means there will be 1 Independent feature and 1 Dependent Feature.

In multi Linear regression, there will be Multiple Independent feature and 1 Dependent Feature.

ONE Problem Statement :-

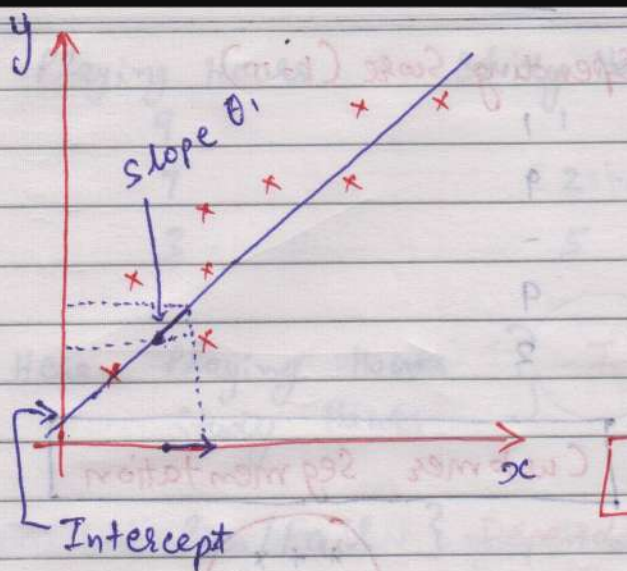
DATASET

Height	Weight

AIM :- To Create a model

I/P - Height

Predict - Weight



Equation of a Straight Line

$$y = mx + c$$

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

$$h_\theta(x) = \theta_0 + \theta_1 x_1$$

Intercept :- θ_0 :- In the best fit line, when the x value is 0, at that time where is the best fit line meating the y axis, that particular point is intercept.

Slope $\theta_1 x_1$:- With the unit movement in the x axis, what is the unit movement in y axis.

COST FUNCTION

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

Mean Squared Error.

$h_\theta(x)^{(i)}$: predicted datapoint

$y^{(i)}$: Actual datapoint

where m = No. of datapoints.

$h_\theta(x)^{(i)}$ = Predicted values

$(y)^{(i)}$ = Actual values.

Final Aim :-

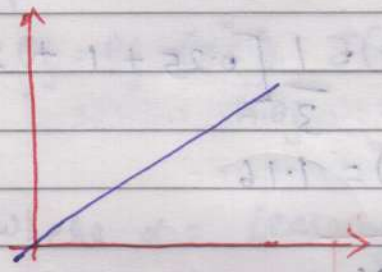
Minimize.

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

Where θ_0 = Intercept
θ_1 = Slope.

$\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2$ = MSE
Mean Squared Error.

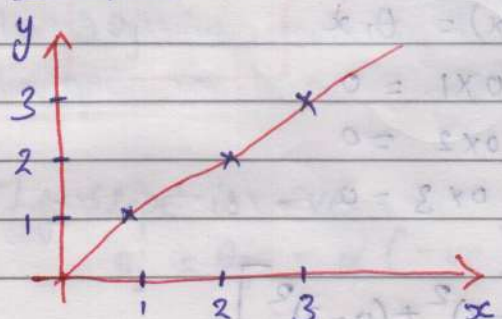
Let us consider $\theta_0 = 0$



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

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

Example :-



$\theta_1 = 1 \quad \theta_0 = 0$

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

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

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

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

Training dataset.

x	y
1	1
2	2
3	3

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

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

$$= 0$$

Here when $\theta_1 = 1 \quad J(\theta_1) = 0$

Example 2:- If $\theta_1 = 0.5$ $\quad h_\theta(x) = \theta_1 x$

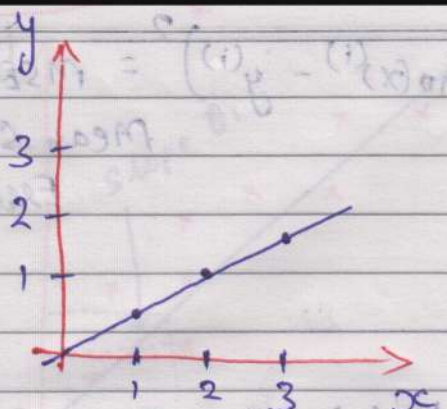
$h_\theta(x) = 0.5 \quad x = 1$

$h_\theta(x) = 1 \times 0.5 \quad x = 2$

$= 1$

$h_\theta(3) = 3 \times 0.5 \quad x = 3$

$= 1.5$



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

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

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

$$J(\theta_1) = 1.16$$

When $\theta_1 = 0.5$ $J(\theta_1) = 1.16$

If $\theta_1 = 0$ $h(\theta)(x) = \theta_1 x$

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

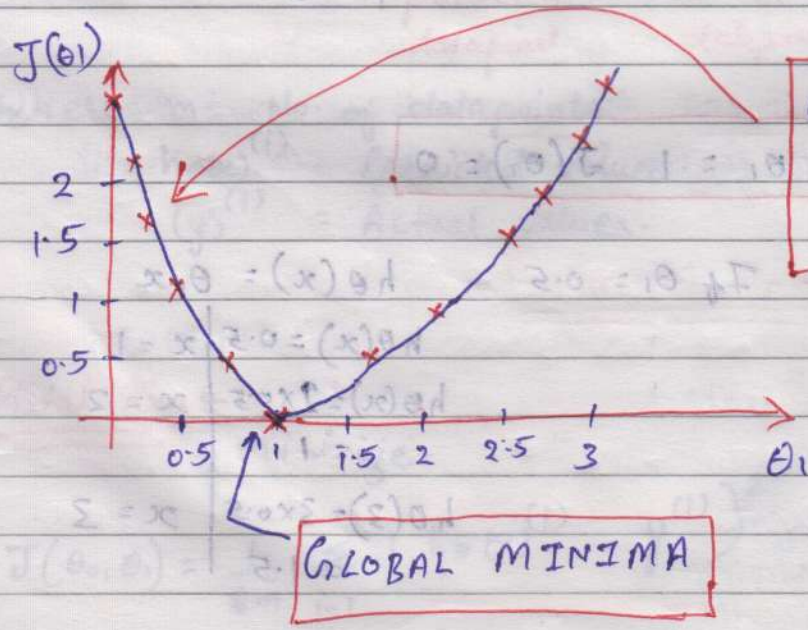
$x=2$ $h_\theta(2) = 0 \times 2 = 0$

$x=3$ $h_\theta(3) = 0 \times 3 = 0$

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

$$= 4.66$$

When $\theta_1 = 0$ $J(\theta_1) = 4.66$



Convergence Algorithm { optimize the changes of θ_i value }

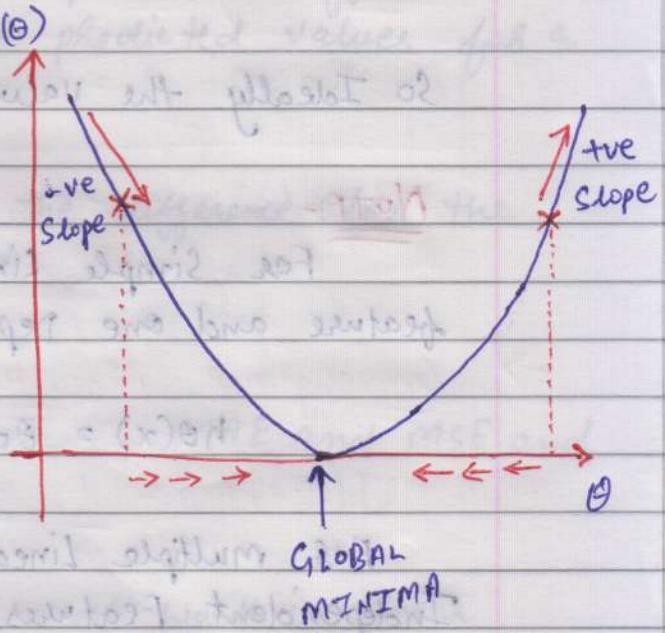
Repeat until Convergence

{

$$\theta_j = \theta_j - (\alpha) \frac{\partial}{\partial \theta_j} J(\theta_j)$$

where α = Learning Rate.

$$\frac{\partial}{\partial \theta_j} J(\theta_j) = \text{Slope}.$$



If Slope is -ve

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

$\therefore$ θ_j will go towards Global Minima. (+ve Direction)

If Slope is +ve

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

$\therefore$ θ_j will go towards Global Minima (-ve Direction)

Question:- What is the Learning Rate?

Answer:- Learning Rate decides the speed of convergence. { How Fast you want to converge }

If Learning Rate (α) = 0.0000001 means if α is very small then it will take more time to come towards Global Minima.

If α value is very large, then it will jump here and there very fast, and would not come near Global minima.

So Ideally the value of $\alpha = 0.001$

NOTE:-

For Simple Linear Regression; where 1 Independent feature and one Dependent Feature.

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

For multiple Linear Regression, where multiple Independent Features {Say 3} and one Dependent Feature

$$h\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

Questions:-

(1) What is Linear Regression?

Ans: It is a statistical method that is used for predictive analysis. Linear regression algorithm shows a linear relationship between a dependent and one or more independent variables, hence called as linear regression.

(2) How we can calculate Error in linear Regression?

Ans:
- Error is the difference between predicted value and real value.
- Linear regression most often uses mean-squared error (MSE) to calculate the error of the model