Data Science & AI

Machine Learning

Supervised Learning

Lecture No.- 01

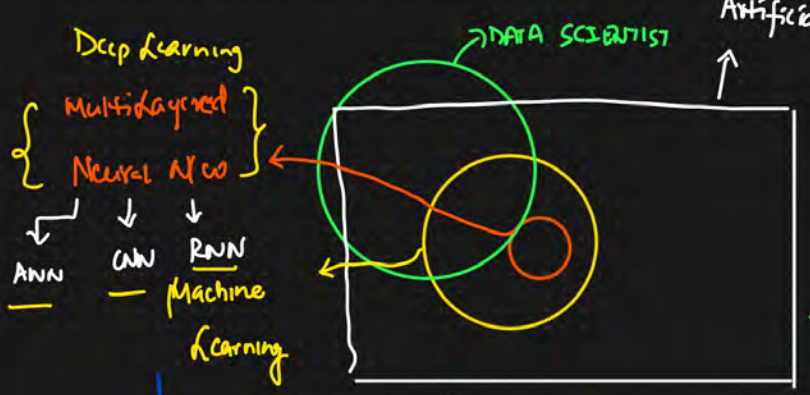
By- Krish Naik Sir

Topics to be Covered

- Topic — Introduction
- Topic — Regression & Its types
- Topic
- Topic
- Topic

Machine Learning.

AI vs ML Vs DL Vs DATA SCIENCE



AI Application

↓

ChatGpt

Google BARD

Artificial Intelligence

↳ Smart applications that can perform its own task without any human intervention

AI App

→ Netflix → Movie Recommendation

Self Driving Cars → Computer Vision Model, Image processing

↳ Statstools → Analyze, Visualize, Prediction, forecasting. Hypothesis Testing

Machine Learning And Deep Learning

- Supervised ML
  - → Regression ✓
  - → Classification ✓
- Unsupervised ML
  - → Clustering ⇓ GROUPS
- Reinforcement Learning
- Semi Supervised ML {Generative AI}

I/p → [Model] → O/p Answer

House Price Prediction [Bangalore] [Supervised]

I/p

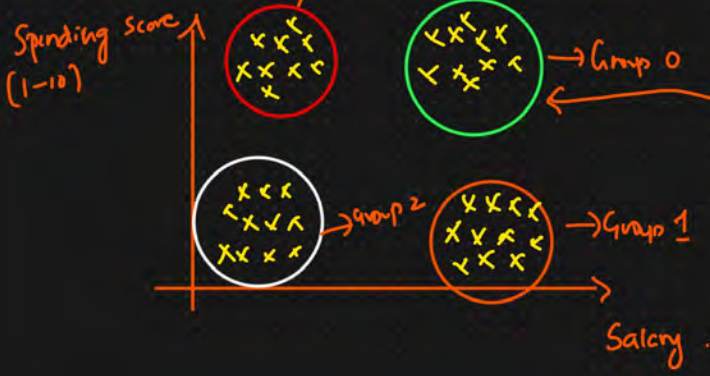
House size	No. of Rooms	Location	O/p. Price [cr] [Continuous Value] ⇓ [Regression] ←
→ 2500 sqft	4	Mumbai	2 cr
→ 1800 sqft	3.5 bhk	Bangalore	1.5 cr

CLASSIFICATION ←

O/p Fixed Categories.

Binary or Multiclass Categories.

② Unsupervised → Clustering



Salary	Spending Score

Iphone 16 ← 100 | 20%

↓ kW

2%

Clustering Similar Data Points

Supervised ML

Regression

① Linear Regression
② Ridge, Lasso, ElasticNet
③ Decision tree
④ KNN, Naive Bayes
⑤ Random Forest
AdaBoost
Xg boost
SVM

Classification

① Logistic Regression

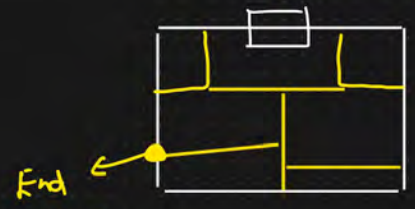
Unsupervised ML

① PCA (Principal Component Analysis)
② K Mean Clustering
③ Hierarchal Clustering
④ DBScan clustering.
⑤ t-SNE

ANN, RNN, CNN

③ Reinforcement learning

↓Rewards ↑↑↑

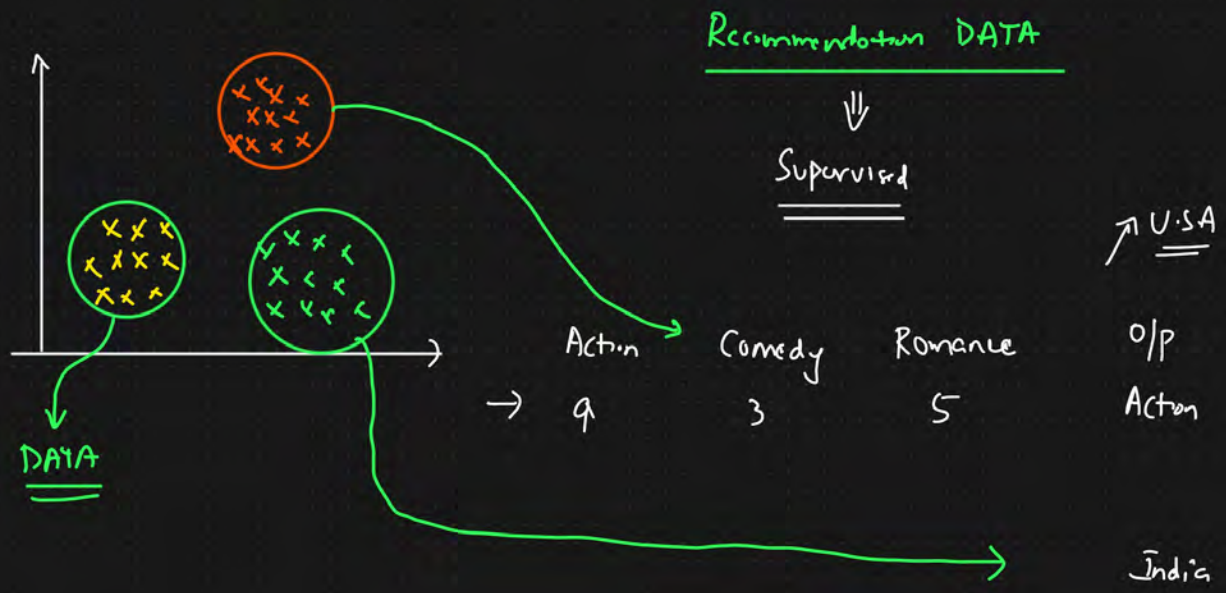


Environment

GTA

④ Semi Supervised ML

Partially Supervised + Unsupervised



Generative AI → LLM, LIM [Large Image Models]

↓

Large Language Models

DATA Increase

500 page

↓

5 months

CAT

Indian

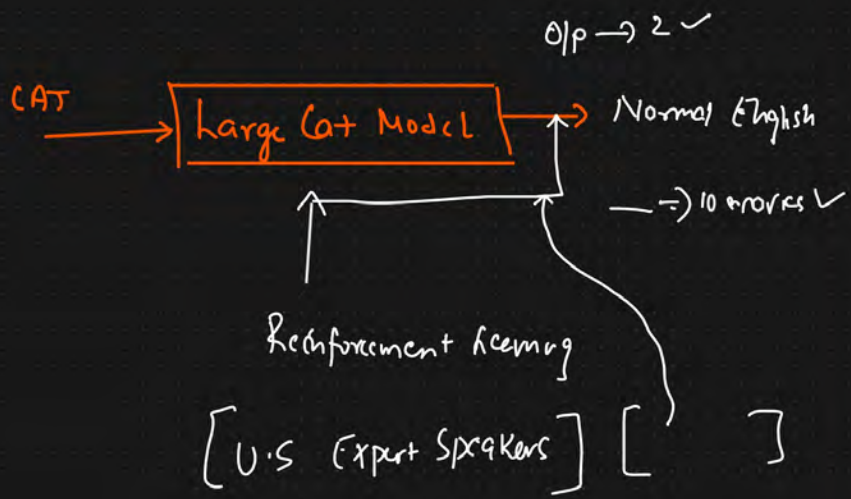
1000 pages

6 months

CAT

↓ US-English

CAT

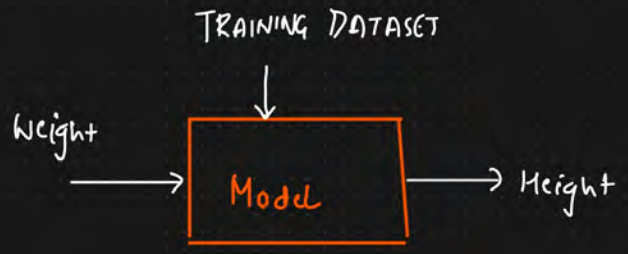


Simple Linear Regression [Supervised ML]

O/P : Continuous Value

Dataset :

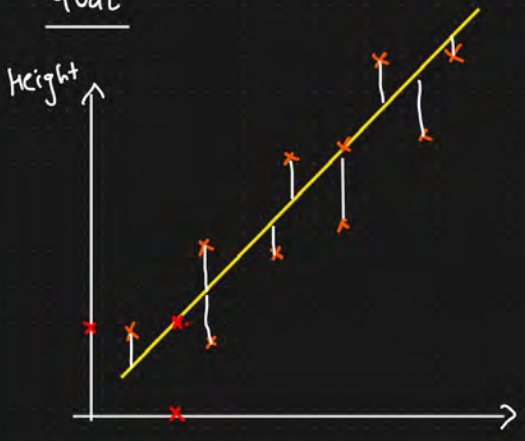
I/P → Weight	O/P → Height
→ 74	170
→ 80	180
→ 75	170
→ -	-
→ -	-

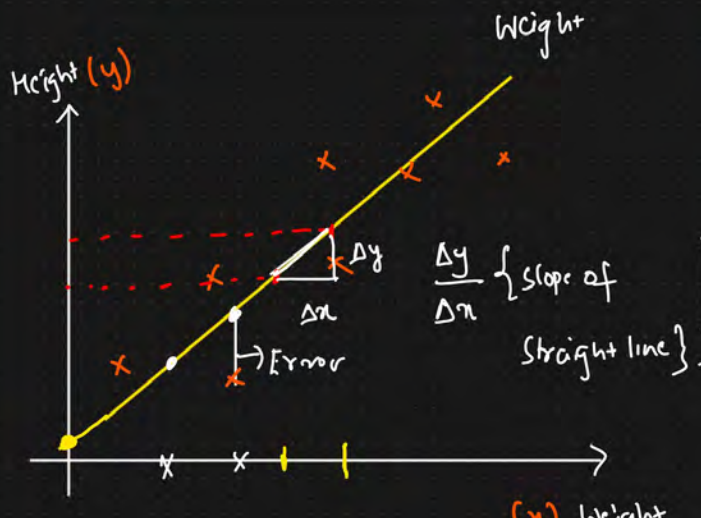


Goal

To create a best fit line in such a way the summation of the error should be minimum

Goal





$\frac{\Delta y}{\Delta x}$ {slope of straight line}.

$\hat{y} = mx + c$

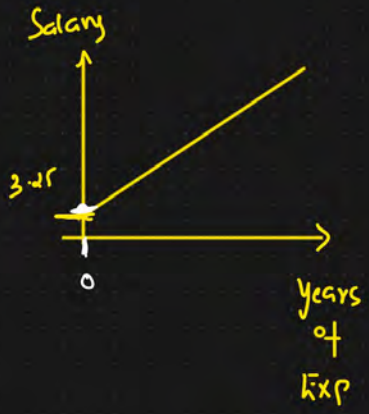
$\hat{y} = \beta_0 + \beta_1 x_1$

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

$\theta_0 \Rightarrow$ Intercept

$\theta_1 \Rightarrow$ Slope or Coefficient.

$x_1 \rightarrow$ Data points.



Cost function — [Error].

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

(y_i: Actual, $h_\theta(x)$: predicted)

[Mean Squared Error].

$\theta_0, \theta_1 \rightarrow$ change Error value

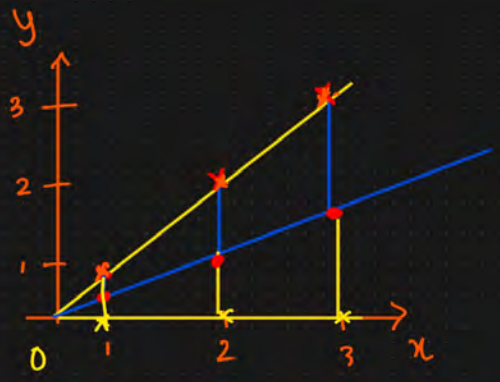
n = no. of datapoints

$y_i \Rightarrow$ Actual value

$h_\theta(x) \Rightarrow$ predicted value.

Final Goal : [In order to get the best fit line]

$$\text{Minimize}_{\theta_0,\theta_1} \; J(\theta_0,\theta_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - h_\theta(x)_i\right)^2$$



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

Lets $\theta_0 = 0$

$\boxed{h_\theta(x) = \theta_1 x_1}$

Dataset

X	y
1	1
2	2
3	3

$h_\theta(x) = \theta_1 x_1$	$\theta_1 = 1$	$\theta_1 = 0.5$	Let $\theta_1 = 0$
$x = 1$	$h_\theta(x) = 1$	$h_\theta(x) = 0.5$	$h_\theta(x) = 0$
$x = 2$	$h_\theta(x) = 2$	$h_\theta(x) = 1$	$h_\theta(x) = 0$
$x = 3$	$h_\theta(x) = 3$	$h_\theta(x) = 1.5$	$h_\theta(x) = 0$

$$J(\theta_0,\theta_1) = \frac{1}{n}\sum_{i=1}^{n}(y_i - h_\theta(x)_i)^2$$

$\theta_1 = 1$

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

$J(\theta_1) = 0$

$\theta_1 = 0.5$

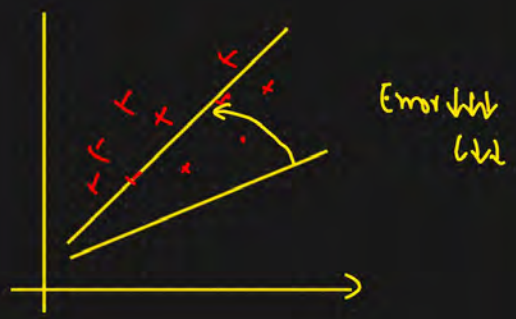
$$J(\theta_1) = \frac{1}{3}\left[(1-0.5)^2 + (2-1)^2 + (3-1.5)^2\right].$$

$= 1.16$

$\theta_1 = 0$

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

$J(\theta_1) = 4.66$



Convergence Algorithm {Optimize the change of θ_0, θ_1 to Global Minima}.

Repeat until convergence

$\alpha = 1$

{

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

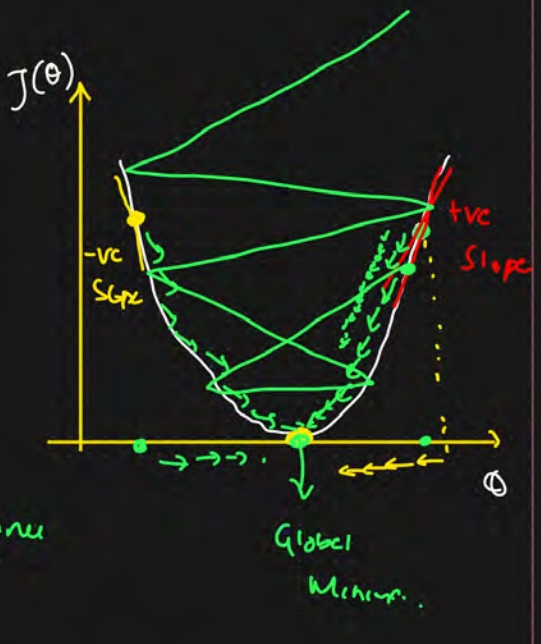
}

$\Rightarrow$ Derivative of slope.

α $\rightarrow$ Learning Rate

α = Speed of convergence

$\alpha \Rightarrow$ Huge value.



Scenario 1

① $\theta_1 = \theta_1 - \alpha\,(+ve)$

$\theta_{1new} = \theta_{1old} - (+ve)$

$\theta_{1old} >> \theta_{1new}$.

Scenario 2

① $\theta_1 = \theta_{1old} - \alpha\,(-ve)$

$= \theta_{1old} + (+ve)$

$\theta_{1old} << \theta_{1new}$

Conclusion

Repeat until Convergence

{

$\theta_j := \theta_j - \alpha \dfrac{\partial}{\partial \theta_j} J(\theta_1)$.

}

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

⇓

Mean Square Error

THANK - YOU