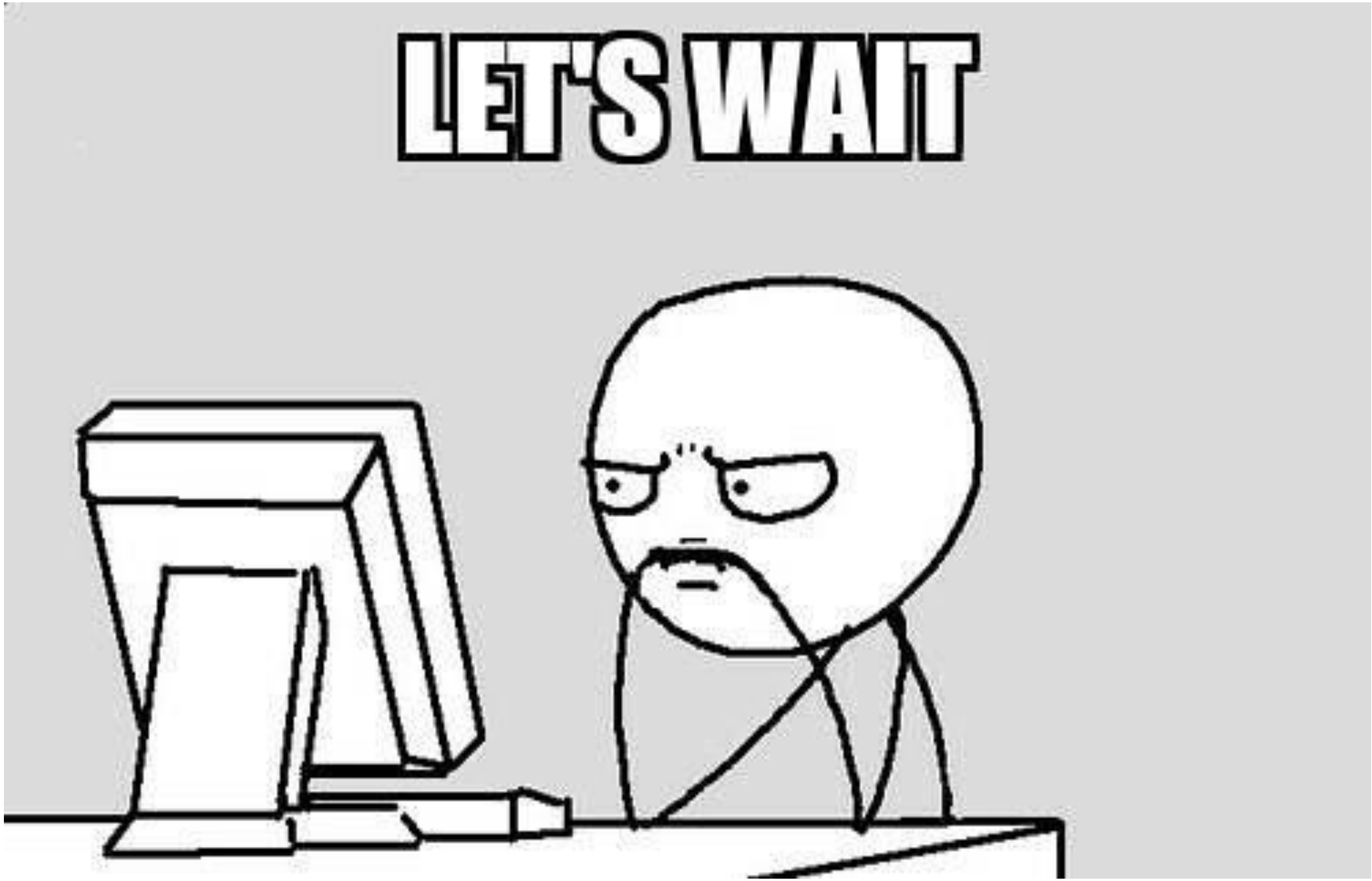

LET'S WAIT

#LifeKoKaroLift

Data Science Certification Program

→ Linear Algebra (Matrics):-

$n \rightarrow$ data points

$m \rightarrow$ features

$y = mx + c$ [SLR]

* L.R

$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_0$ [MLR] ✓

Matrix Representⁿ of MLR

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_m \end{bmatrix} \begin{bmatrix} X_{11} & X_{21} & \cdots & X_{m1} \\ X_{12} & X_{22} & \cdots & \\ \vdots & & \ddots & X_{mn} \end{bmatrix} + \begin{bmatrix} \beta_0 \\ \beta_0 \\ \vdots \\ \beta_0 \end{bmatrix}$$

→ **Basic Equation of NN :-**

$W^T X^T + b$

Weights (Eg. slope)

Output

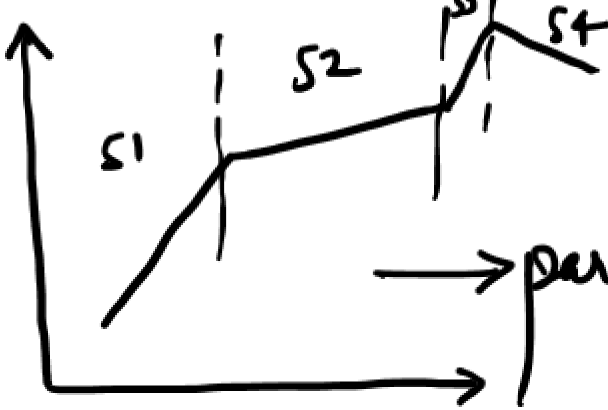
$$y = f(Wx + b)$$

bias (Eg. intercept)

features

Activation f^n

[Lin Reg, Log.R]

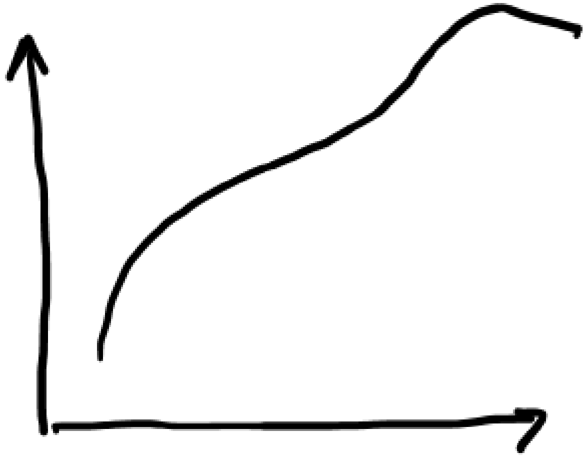


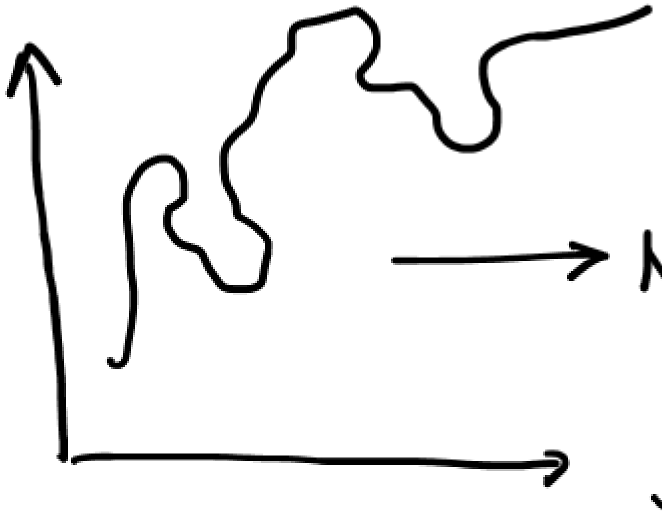
[Tree based]

→ partially linear

→ Non-linear

→ SVM(radial)





→ NN

↓

Universal f^n approximator

→ Data

→ Computatⁿ

→ Non-interpretability

Almost Solved

fin-tech / Telecom } Regulatory

RBI

TRAI

$$\begin{bmatrix} 2 & 5 \\ 8 & 0 \end{bmatrix}_{2 \times 2} \qquad \begin{bmatrix} 5 & 1 & 3 \\ 6 & 8 & 2 \end{bmatrix}_{2 \times 3}$$

$$2 \times \boxed{2 \times 2} \times 3$$

$$\downarrow$$

$$\boxed{2 \times 3}$$

$$= \begin{bmatrix} 2 \times 5 + 5 \times 6 & 2 \times 1 + 5 \times 8 & 2 \times 3 + 5 \times 2 \\ 8 \times 5 + 0 \times 6 & 8 \times 1 + 0 \times 8 & 8 \times 3 + 0 \times 2 \end{bmatrix}$$

$$= \begin{bmatrix} 40 & 42 & 16 \\ 40 & 8 & 24 \end{bmatrix} \quad \text{Ans}$$

Ques:-

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 8 & 9 \end{bmatrix}_{3\times 2} \qquad B = \begin{bmatrix} 2 & 0 \\ 8 & 6 \end{bmatrix}_{2\times 2}$$

3×2

find dot product of A & B.

Sol

$$\begin{bmatrix} 3\times2+2\times8 & 3\times0+2\times6 \\ 1\times2+2\times8 & 1\times0+2\times6 \\ 8\times2+9\times8 & 8\times0+9\times6 \end{bmatrix} = \begin{bmatrix} 22 & 12 \\ 18 & 12 \\ 88 & 54 \end{bmatrix}$$

→ <u>Calculus</u> :-

* Differentiation (Chain rule)
* Complete derivative Vs partial derivatives
* Maxima / Minima

$\longrightarrow \quad y = x^n$

$\frac{dy}{dx}$ [Differentiatⁿ of y w.r.t x]

$= \frac{dx^n}{dx}$

$= n x^{n-1}$

Important

Q: find the differentⁿ of $7x^3 + 10$ w.r.t x

Sol

$y = 7x^3 + 10$

$\frac{dy}{dx} = \frac{d(7x^3+10)}{dx}$

$= \frac{d(7x^3)}{dx} + \cancel{\frac{d(10)}{dx}}$

$= 7\frac{dx^3}{dx} + 0 = \underline{\underline{21x^2}}$

→ Chain rule :- [Back propagation]

$$\frac{dx^n}{dx} = nx^{n-1}$$

$$y = (\underbrace{x^2 + 2x - 3})^4$$

$$\frac{dy}{dx} = 4(\underbrace{x^2 + 2x - 3})^3 (2\underbrace{x} + 2)(1)$$

$$y = (\underbrace{x^2+2x-3}_{\to\, t})^4 \qquad \underbrace{t = x^2+2x-3}$$

$$y = t^4 \Rightarrow \frac{dy}{dt} = 4t^3 \qquad \frac{dt}{dx} = 2x+2$$

$$\frac{dy}{dx} = \frac{dy}{dt} \times \frac{dt}{dx}$$

$$= 4t^3(2x+2)$$

$$= 4(x^2+2x-3)^3(2x+2)$$

Qus:- $y = \left(\underbrace{(x^2+3)^6 + 9x^2 + 4x^3 + 6}\right)^5$

$t = (x^2+3)^6 + 9x^2 + 4x^3 +$

$y = t^5$

$\frac{dy}{dt} = 5t^4$

$t = \gamma^6 + 9x^2 + 4x^3 + 6$

$\gamma = (x^2+3) \Rightarrow \frac{d\gamma}{dx} = 2x$

$$\boxed{\frac{dt}{dx} = 6\gamma^5 \frac{d\gamma}{dx}} + 18x + 12x^2 + 0$$

$$\frac{dy}{dx} = \frac{dy}{dt} \times \frac{dt}{dx}$$

$= 6(x^2+3)^5(2x) + 18x + 12x^2$

$= 5t^4 \times \left[6(x^2+3)^5(2x) + 18x + 12x^2\right]$

$= 5\left((x^2+3)^6 + 9x^2 + 4x^3 + 6\right)^4 \left[6(x^2+3)^5(2x) + 18x + 12x^2\right]$

→ Complete Derivatives

* $\frac{dy}{dx}$

* $t = r^6 + 9x^2 + 4x^3 + 6$

$$\frac{dt}{dx} = \underbrace{6r^5 \frac{dr}{dx} + 18x + 12x^2}$$

Partial derivatives (NN) | ML

* $\frac{\partial y}{\partial x}$ $\partial \rightarrow$ dow

* $t = r^6 + 9x^2 + 4x^3 + 6$

$$\frac{\partial t}{\partial x} = 0 + \underbrace{18x + 12x^2}$$

- It assumes other than x everything is constant } holds true

→ **Application :-**

$J \rightarrow$ cost f^n

$$\left. m = m - \alpha \underbrace{\frac{\partial J}{\partial m}} \right]_{c = \text{constant}}$$

$$\left. c = c - \alpha \frac{\partial J}{\partial c} \right]_{m = \text{constant}}$$

→ Cost f^n :- (Regression) [MSE - Mean Square Error

Imp

$$J = \frac{1}{N}\sum_{i=1}^{N}\left[y_{a_i} - (mx_i + c)\right]^2$$

[Parabola Eq^n

Mean Square Error :-

Sum / N

$$J = \frac{1}{N}\sum_{i=1}^{N}\left[y_i - (mx_i + c)\right]^2$$

→ Update Eq^n of m :-

$$m' = m - \cancel{\alpha \frac{\partial J}{\partial m}} \to 0$$

$$\frac{\partial J}{\partial m} = \frac{\partial}{\partial m}\left\{ \frac{1}{N} \sum \left[y-(mx+c)\right]^2 \right\}$$

$$= \frac{1}{N} \sum 2 \cdot \left[y-(mx+c)\right]^{2-1} \left(0-(x+0)\right)$$

$$= \frac{2}{N} \sum \left[y-(mx+c)\right](-x) = \frac{-2}{N} \sum \left[y-(mx+c)\right](x)$$

→ Update Eq^n of c :-

$$c' = c - \cancelto{0}{\alpha \frac{\partial J}{\partial c}}$$

$$\frac{\partial J}{\partial c} = \frac{\partial}{\partial c}\left\{ \frac{1}{N} \sum \left[y - (mx + c) \right]^2 \right\}$$

$$= \frac{1}{N} \sum 2 \cdot \left[y - (mx + c) \right] \left(0 - (0 + 1) \right)$$

$$= \frac{-2}{N} \sum \left[y - (mx + c) \right]$$

$\rightarrow$ Maxima / Minima :-

$$\left\{ y = \underline{\underline{x^2 + 4}} \right\}$$

Max | min

first derivative

Second derivative

$$\Rightarrow \underbrace{\frac{dy}{dx} = 0}$$

$$2x = 0$$

$$\boxed{x = 0}$$

$$y \rightarrow \text{max} \mid \text{min} \ @ \ x = 0$$

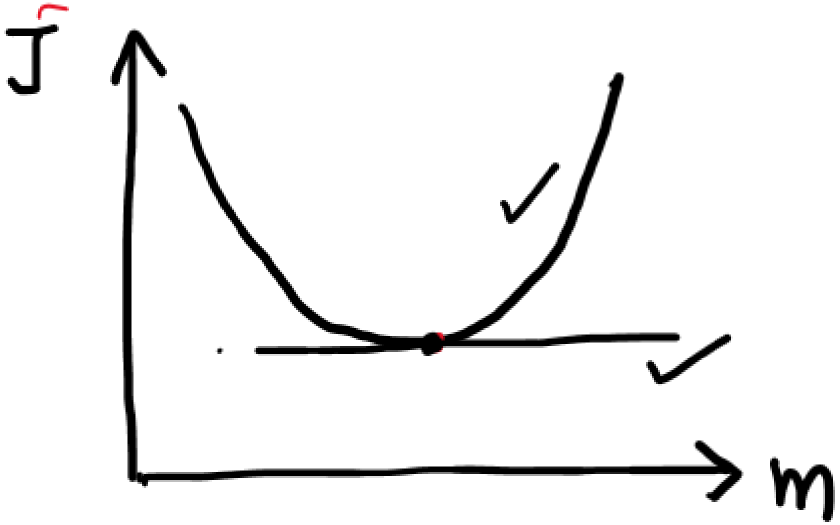
$$\Rightarrow \frac{d^2y}{dx^2} > 0 \rightarrow \text{Minima}$$

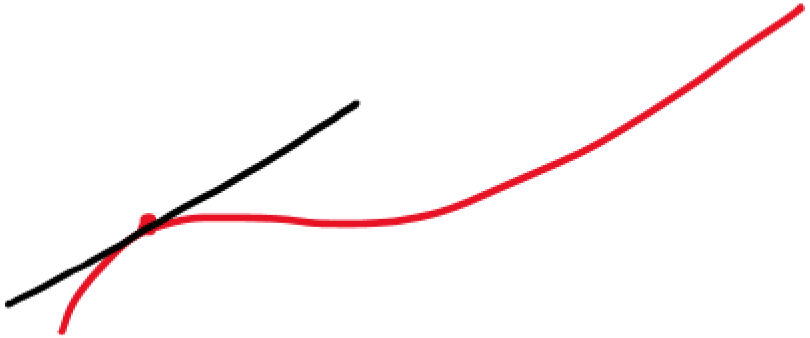
$$\frac{d^2y}{dx^2} < 0 \rightarrow \text{Maxima}$$

$$\underbrace{\frac{d^2y}{dx^2} = 2 > 0}$$

* In gradient descent (LR), how algorithm knows minima is achieved.

Slope of curve = $\frac{\partial J}{\partial m}$ [gradient]





At minima $\frac{\partial J}{\partial m} = 0$ [gradient becomes zero]

$m' = m - \alpha \cancel{\frac{\partial J}{\partial m}} \to 0$

$c' = c - \alpha \cancel{\frac{\partial J}{\partial c}} \to 0$