Perceptron ⇒

① SLP

② MLP

input

Input

x_1

w

b_1

x_2

w_2

x_3

w_3

Hidden

Prior knowledge

output L

$y =$

$\sum x_i w_i$ $At(y)$

Sigmoid, Tanh, Soft-max, relu

$\hat{y}$

$$\sum x_i w_i = [x_1 w_1 + x_2 w_2 + x_3 w_1] + b_1$$

Forward prop

Ex ⇒   School ⇒ Input Clas ⇒   Attention → Notes ⇒   Output

(x_i) Weights   Bias   Actual
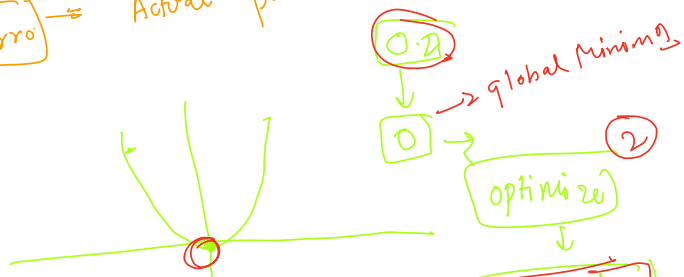   ↓   संगठन
                    0   ↓   ↓
                                              Predicted

Backward propagation

① Actual

② → Predicted

Error = 5·2 − 5 = 0·2
↓
Actual − predicted = different

Loss function — Error =

①

0·2

0 → global Minima

②

optimize

Weigent- Update
↓
Error Reduction → optimization

Optimizato

1) Gradient descent
2)

Summary → x_i^a , Weight

3) Activation f
4) Bias
5)

} — Forward popagati

2) Loss fut
Optimation
Updc weign

} → Backward propagation.

→ Weight update → Process → < Weight updation formula
Chain Rule of diffentiation

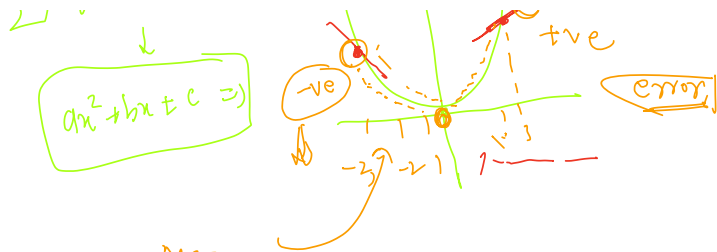→ Method → Activation →, Sigmoid
→ Tan ħ
→ Relu
→ Leaky Rely

① Weight updation formula →

Learning Rate

$[W]_{now} = W_{old} - \eta \frac{\partial L}{\partial W_{old}}$ → 1) MSE
2) MAE
3)

Slope

$\frac{1}{n}(y-\hat{y})^2$ → → upward

$ax^2 + bx + c \Rightarrow$

$-ve$ $+ve$

error

$-3 \quad -2 \quad 1 \quad 1$

MSE

$$\Rightarrow \boxed{W_{new} >> W_{old}}$$

① $\Rightarrow)$ $\boxed{W_{new}} = W_{old} - \eta(-ve)$

      $W_{od} + \,^{ve} =$
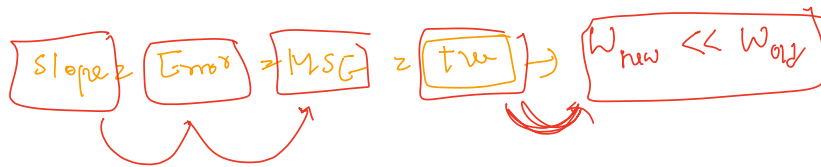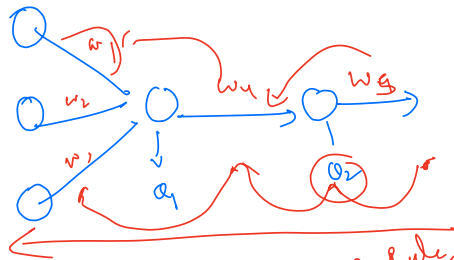
② $\boxed{Slope} \in \boxed{Error} \Rightarrow \boxed{MSG} = \boxed{+ve} \Rightarrow \boxed{W_{new} << W_{old}}$

$\boxed{-ve} \Rightarrow बड़ा होना चाहिये = New \, W$

$\boxed{+ve} = Small$

③ $\underline{Chain \, Rule} \Rightarrow$

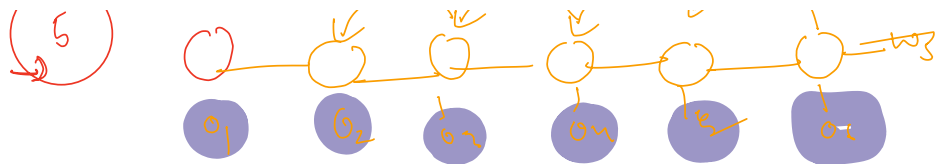$\downarrow$

w_1
w_2 w_4 w_5
w_3 O_1 O_2

$\boxed{W_{new} = W_{old} - \eta \dfrac{\partial L}{\partial W_{old}}}$ $\rightarrow$ chain Rule

$\dfrac{\partial L}{\partial O_2} \times \dfrac{\partial O_2}{\partial w_4} \rightarrow$ Partiv derivd

$\downarrow$

Chein Rule

①

Diffetention => 0.028 —> 0.024 ···· ···· 0.0001

weight

$W_{new} = (W_{old}) - m (0.001)$

Small

=> 2gs intimli'

Vanishing Gradient

$W_{new} \approx W_{old}$

1) Activation funcion

2)

Sigmoid => 0.8 0 1

Tanh = -1 1 =>

(6) Activation Function → Non Linearity

→ Non linearity
→ Hidden
→ Pattern & Hidds

← $y = mx + b$ = Linearity

Sigmod = $\frac{1}{1+\theta^{-x}}$ = Non linear

Input → Funtion 2.5 3.1 1.8

(1)
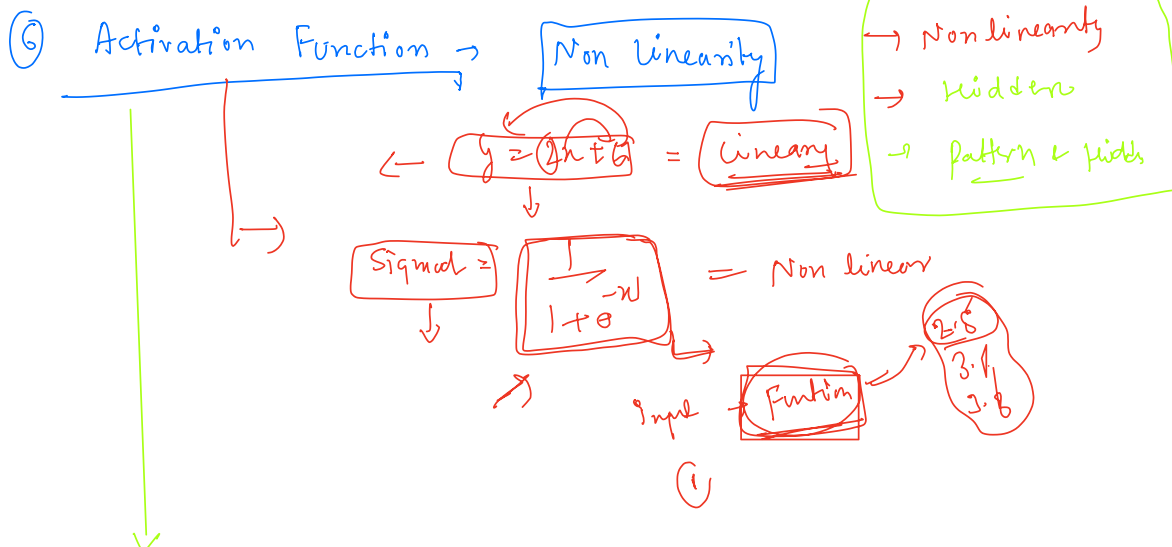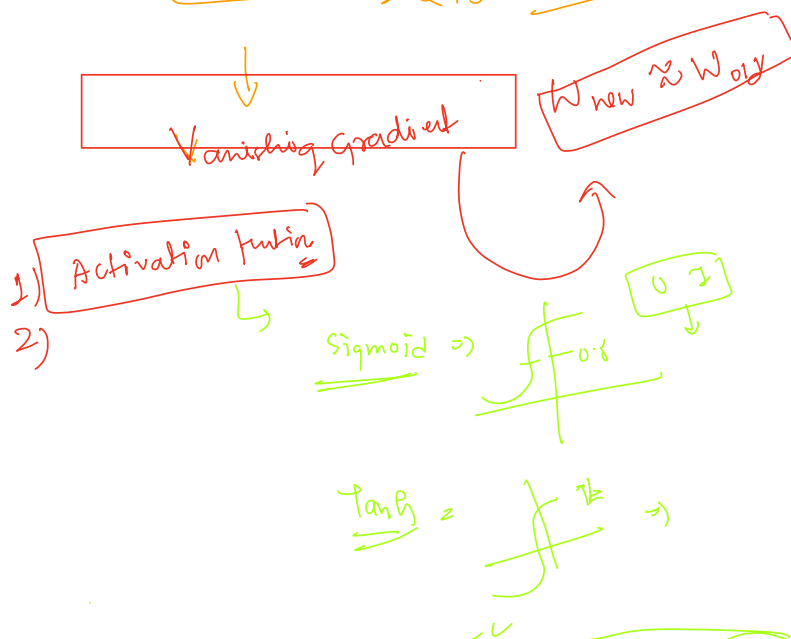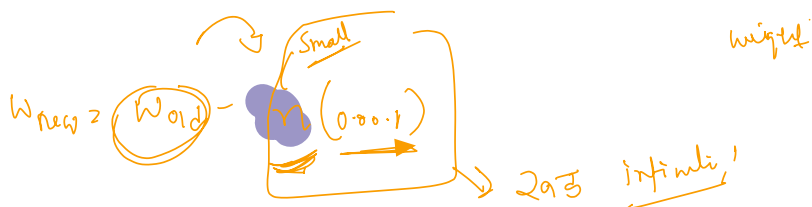
Types ⟹ Sigmoid , Tan h , Softmax , Relu , leakyRelu

① Sigmoid ⟹ Non Linearity ⟹

i) $\dfrac{1}{1+e^{-x}}$

②

0.25 (σ w)

0

(1) ⟹ Not a 0 Centric ⟹ Complex

2)