

## *Gradient descent*

*Step – 1: Inputs will pass through weights and  
the activation of linear combinations is provided by a Neuron*

### *Forward prpogation*

$$\text{linear combinations} = b + W * X$$

$$\text{Activation on linear combinations} = A(b + W * X)$$

*Step – 2: Will caculate the error*

*Step – 3: will minmise the error by going back and updating the weights*

### *Backward prpogation*

*Forward prpogation: Inputs*

*Backward prpogation: Updae the weights*

*It is a iterative process untill we minise the error*

*By using Gradient descent algorithm will update the weights*

$$w_{new} = w_{old} - \alpha * \frac{dJ}{dw_{at=w_{old}}}$$

$$X = [1, 2, 3, 4, 5]$$

$$Y = [11, 12, 13, 14, 15]$$

$$w_{new} = w_{old} - \alpha * \frac{dJ}{dw_{at=w_{old}}} \quad (\text{we are applying on LR})$$

Step – 1: wil intilaise weights and bias randomly

Step – 2: inputs pass through weights and will get output ( $y_{predictions}$ ) =  $b + w * x$

Step – 3: Calculate the error<sup>2</sup> =  $(y_a - y_p)^2$

Step – 4: Calculate MSE or Cost function  $(J) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_a - y_p)^2$

Step – 5:  $\frac{dj}{dw} = w_d = 2 (y_a - y_p) (-x)$

Step – 6: learning rate = 0.1

Step – 7:  $w_{new} = w_{old} - \alpha * w_d$

step – 8: All above repeated 1000 times  $J \cong 0$   $w_{new} = w_{old}$

**EWA NPTEL video**

Ch+j+man

5 == 6

$j == > error == > pred == > input$  we == > random

2 days : Read all the word documents nptel

standford uni

9 am to 10 am exam: Descriptive

10am to 1pm: Text preprocessing