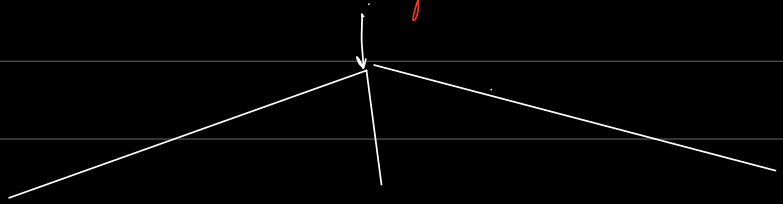
Forward and Backward propagation

Q → How we are training Neural Network ?

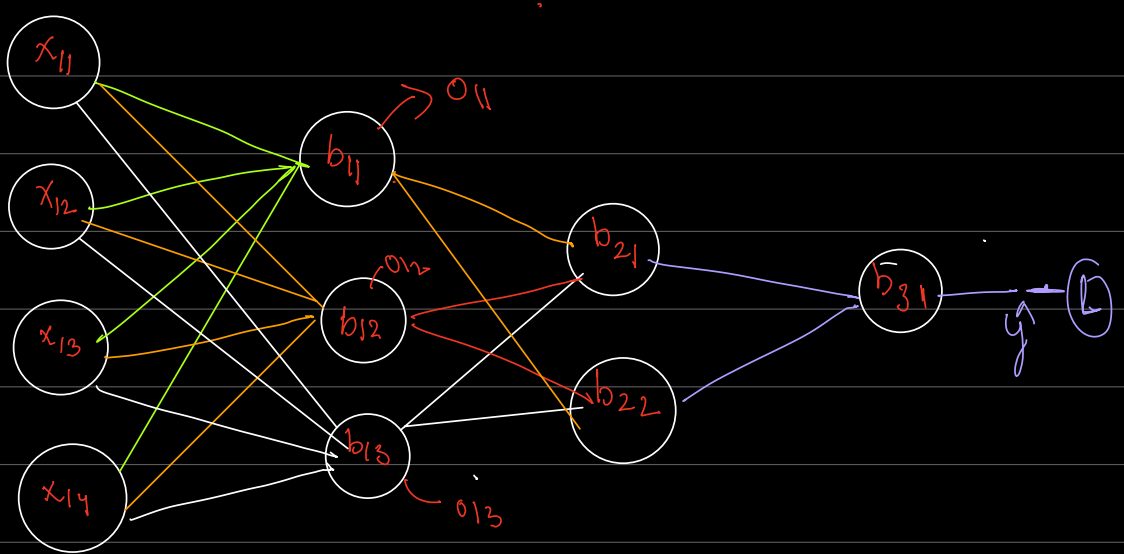


Forward propagation → ①

Loss Function → Concept → ① $(y-\hat{y})^2$

Backward propagation → Mathmatecid

(1) Forward propagation :-



$4 \times 3 = 12$ $\quad 3 \times 2 = 6$ $\quad 2 \times 1$

Forward propagation → 1. Equation

↓ 2 → Matrix → Equation

Eq → $Z = x_1 w_1 + x_2 w_2 + b$

↓

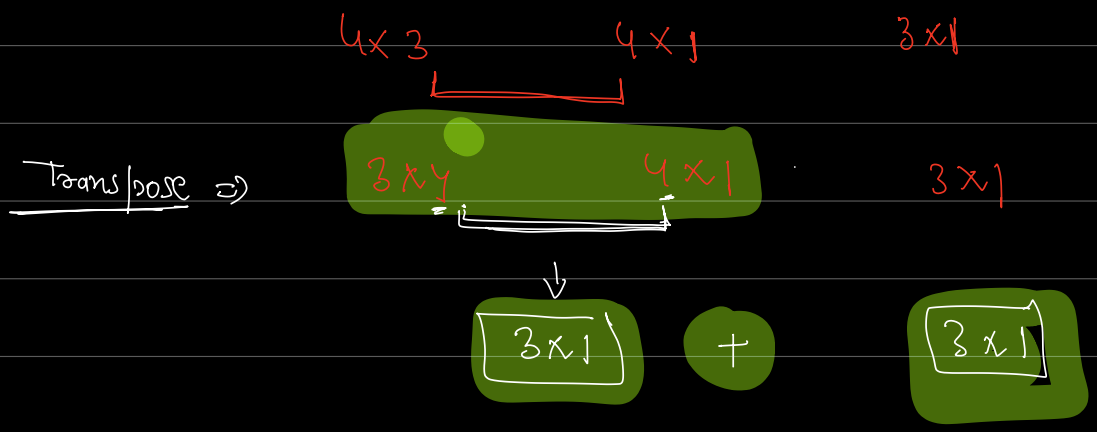
Matrix → equation →

Dataset →	A	B	C	D	P/F
Row 1 →	70	60	30	20	0
Row 2	30	20	50	10	0

Matrix :

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{34} \\ w_{41} & w_{42} & w_{43} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Transpose → Row converted colonom



$$\begin{bmatrix} w_{11}x_{11} + w_{21}x_{12} + w_{31}x_{13} + w_{41}x_{14} \\ w_{12}x_{11} + w_{22}x_{12} + w_{32}x_{13} + w_{42}x_{14} \\ w_{13}x_{11} + w_{23}x_{12} + w_{33}x_{13} + w_{43}x_{14} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Step 1 :

$$z = x_1 w_1 + x_2 w_2 + b_1$$

Step 2 = Multiply with Activation function.

6 [ ]

① Step3 = output = $\hat{y} = \begin{bmatrix} o_{11} \\ o_{12} \\ o_{13} \end{bmatrix}$

S Step2 :

$$\begin{bmatrix} w^2_{11} & w^2_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix}^T \begin{bmatrix} o_{11} \\ o_{12} \\ o_{13} \end{bmatrix} + \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix}$$

$3\times2 \quad\quad 3\times1 \quad + \quad 2\times1$

2×1

$2\times3 \quad\quad 3\times1$

$2\times1 \quad + 2+1$

$$6 \begin{bmatrix} w_{11}o_{11} + w_{21}o_{12} + w_{31}o_{13} \\ w_{12}o_{11} + w_{22}o_{12} + w_{32}o_{13} \end{bmatrix} + \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix}$$

$$\Rightarrow \left\{ = \begin{bmatrix} o_{21} \\ o_{22} \end{bmatrix} \right\}$$

Step → Layer 2 ⇒

$$\begin{bmatrix} w^3_{11} \\ w^3_{21} \end{bmatrix}^T \begin{bmatrix} o_{21} \\ o_{22} \end{bmatrix} + \begin{bmatrix} b_{31} \end{bmatrix}$$

$2\times1 \quad , \quad 2\times1 \qquad \boxed{1\times1}$

$\times$

$1\times2 \qquad 2\times1$

↓

$\boxed{1\times1 \quad + \; 1\times1}$

↓

$$\rightarrow \sigma\left(\left[w_{11}o_{21} + w_{21}o_{22} + b_{31}\right]\right) =$$

⇓

$\boxed{\hat{y}}$ → predicated

↓ ↓

$\hat{y} \rightarrow \boxed{50} \Rightarrow \boxed{10}$

$y \rightarrow \boxed{60}$

difference = Error

Loss function

① Forward propogation → $\hat{y}$ (.....)

② Calculate error using loss function = Error

Error → Resolved → Train our Neural Network

Backpropogation →

It is a algorithm that we ~~are using to~~ train Neural Network

Backpropogation → Req → Random weight

Adjust → weight, bias

Train

Actual weight ← Repeat → Error

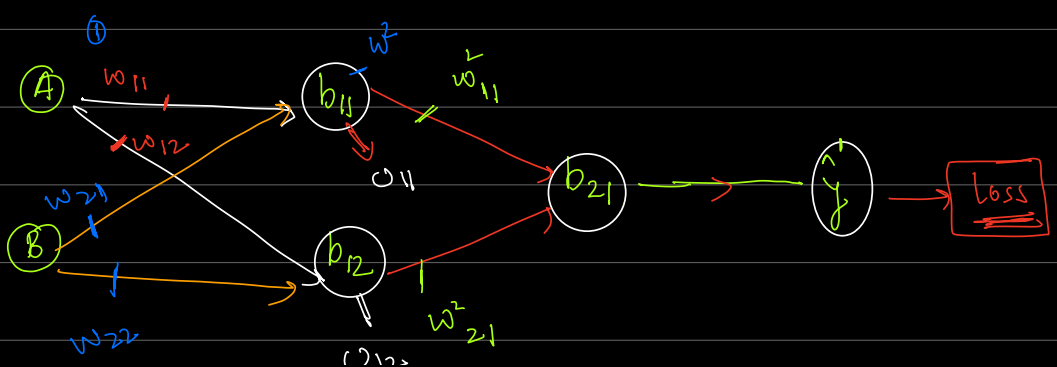
② Actual Bias → Loss → Minimum

Error = Minimum

Algorithm **Backpropogation** ⇒ weight / Bias

1) Forward prop.
2) Gradient Descent

* purpose = Adjust **weight & bias**



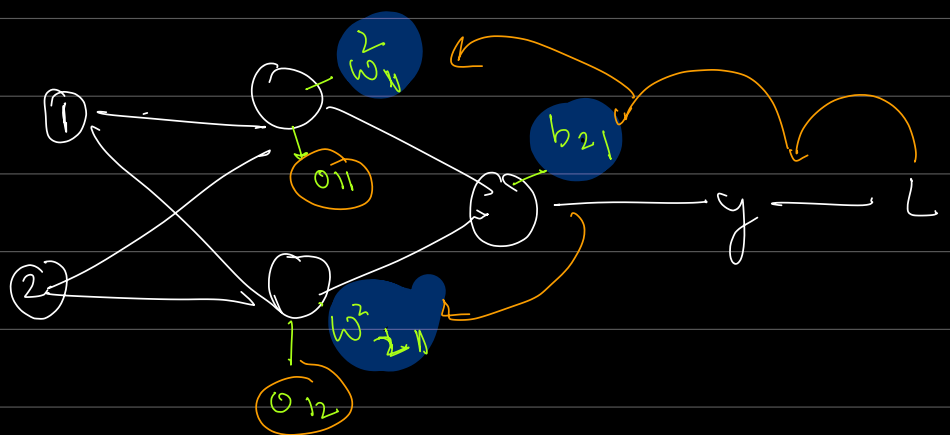
① → Forward

② loss calculation

← Back propagation

↓

Weight



$\hat{y} \Rightarrow z \Rightarrow o_{12} \times w_{21} + o_{11} \times w_{11} + b_{21}$

w_{21} w_{11} b_{21}

Weight update formula ⇒

weight ①

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_{old}}$$

Bias ⇒

$$b_{new} = b_{old} - \eta \frac{\partial L}{\partial b_{old}}$$

w_{21} ⇒

$$w_{21\,new} = w_{21\,old} - \eta \frac{\partial L}{\partial w_{21\,old}}$$

w_{11} ⇒

$$w_{11\,new} = w_{11\,old} - \eta \frac{\partial L}{\partial w_{11}}$$

b_{21} ⇒

$$b_{21\,new} = b_{21\,old} - \eta \frac{\partial L}{\partial b_2} \cdot \frac{\partial b_2}{\partial b_{old}}$$

(0.1)

difference ⇒ $\frac{dy}{dx}$ $\frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2}$

what doing

$\frac{\partial L}{\partial} \Rightarrow \frac{\partial L}{\partial} \times \frac{\partial}{\partial}$ ① ②

$$\frac{\partial L}{\partial w_{11}} \quad \frac{\partial \hat{y}}{\partial w_{11}}$$

$L \rightarrow \hat{y} \rightarrow w_{11}$,

$L \rightarrow \hat{y} \rightarrow w_{21}$ — ②

$L \rightarrow \hat{y} \rightarrow b_{21}$ — ③ ⇒

⇒ ① $\frac{\partial L}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}} (y_i - \hat{y}_i)^2 = -2(y - \hat{y})$

Loss

Predi

$z = x_1 w_1 + x_2 w_2 + b$

② $\frac{\partial \hat{y}}{\partial w_{11}} \Rightarrow \frac{\partial}{\partial w_{11}} \left[o_{11} w_{11} + o_{12} w_{21} + b_{21} \right]$

o_{11} — o_{12} — Const = 0

o_{11}

$\frac{\partial L}{\partial w_{11}} \Rightarrow -2(y - \hat{y}) \times o_{11}$

$$w_{21} \Rightarrow \frac{\partial L}{\partial w_{21}} \Rightarrow \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial w_{21}}$$

$$L \rightarrow y \rightarrow w_{21}$$

$$\frac{\partial L}{\partial \hat{y}} \Rightarrow -2(y-\hat{y})$$

$$\frac{\partial \hat{y}}{\partial w_{21}} \Rightarrow \frac{\partial}{\partial w_{21}} \left[o_{11} \times w_{11} + o_{12} \times w_{21} + b_{21} \right]$$

$o_{11} \times w_{11}$ → 0 ; $o_{12} \times w_{21}$ → 1 ; b_{21} → 0

$$o_{12}$$

$$\frac{\partial L}{\partial w_{21}} = -2(y-\hat{y})\, o_{12}$$

$-2(y-\hat{y})$ → L' ; o_{12} → R

$$\frac{\partial L}{\partial b_{21}} \Rightarrow \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial b_{21}}$$

$-2(y-\hat{y})$ $\Rightarrow$ $\Rightarrow$ $\boxed{-2(y-\hat{y})}$