# Deep Learning | EX01

## Feed Forward Neural Networks (FFNN)

**800**

**Computer Vision | Zahra Amini**

Telegram: @zahraamini_ai & Instagram:@zahraamini_ai & LinkedIn: @zahraamini-ai

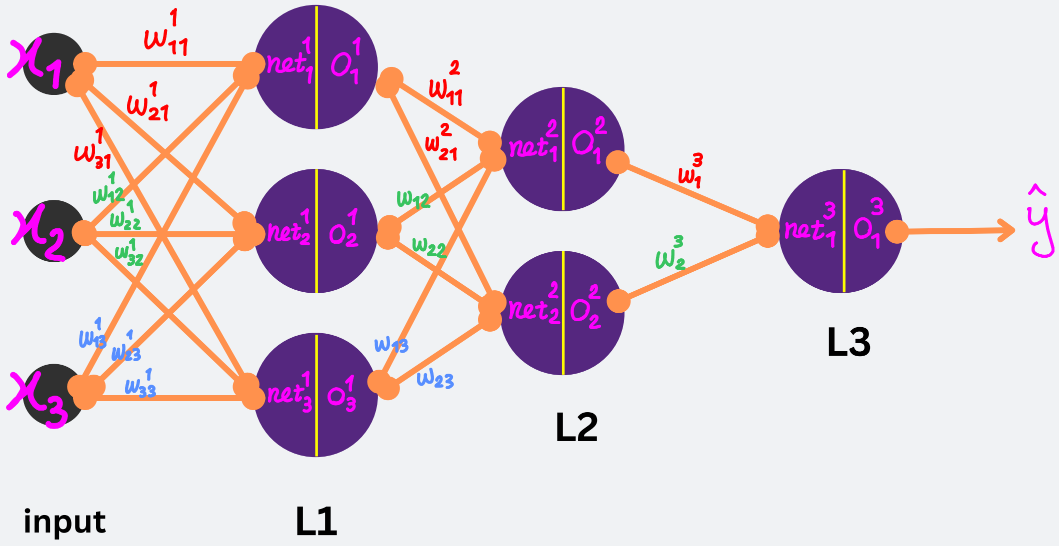
https://zil.ink/zahraamini

1. **For the below feed-forward neural network architecture obtain the following variables:**

$O_1^1 = ?$

$O_2^2 = ?$

$O_1^3 = \text{y\_predicted} = ?$




| inputs | L1-Weights | L1-Biases | L2-Weights | L2-Biases | L3-Weights | L3-Biases |
|---|---|---|---|---|---|---|
| $\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array}\begin{bmatrix} 0.3 \\ 0.5 \\ 0.8 \end{bmatrix}$ | $\begin{bmatrix} \overset{w_{11}}{0.1} & \overset{w_{12}}{0.2} & \overset{w_{13}}{0.2} \\ \overset{w_{21}}{0.4} & \overset{w_{22}}{0.1} & \overset{w_{23}}{0.2} \\ \overset{w_{31}}{0.2} & \overset{w_{32}}{0.3} & \overset{w_{33}}{0.1} \end{bmatrix}$ | $\begin{bmatrix} 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}$ | $\begin{bmatrix} \overset{w_{11}}{0.1} & \overset{w_{12}}{0.2} & \overset{w_{13}}{0.2} \\ \overset{w_{21}}{0.4} & \overset{w_{22}}{0.1} & \overset{w_{23}}{0.2} \end{bmatrix}$ | $\begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix}$ | $\begin{bmatrix} \overset{w_1^3}{0.5} & \overset{w_2^3}{0.1} \end{bmatrix}$ | $\begin{bmatrix} 0.5 \end{bmatrix}$ |

**Activation Functions:**
- L1: Relu
- L2: Tanh
- L3: Sigmoid

: $O_1^1 = ?\ 0.49 \qquad O_2^2 = ?\ 0.621 \qquad \hat{y} = ?\ 0.673$

## Layer #1:

$$net_1^1 = \sum Wx + b = W_{11}^1 \times x_1 + W_{21}^1 \times x_2 + W_{31}^1 \times x_3 + b_1^1 = \left(\overset{0.03}{\cancel{0.1 \times 0.3}}\right) + \left(\overset{0.1}{\cancel{0.2 \times 0.5}}\right) + \left(\overset{0.16}{\cancel{0.2 \times 0.8}}\right) + 0.2 = 0.49$$

$$\begin{bmatrix} net_1^1 \\ net_2^1 \\ net_3^1 \end{bmatrix} = WX + b = \underset{3\times 3}{\begin{bmatrix} 0.1 & 0.2 & 0.2 \\ 0.4 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.1 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} 0.3 \\ 0.5 \\ 0.8 \end{bmatrix}} + \begin{bmatrix} 0.2 \\ 0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.1 \times 0.3 + 0.2 \times 0.5 + 0.2 \times 0.8 \\ 0.4 \times 0.3 + 0.1 \times 0.5 + 0.2 \times 0.8 \\ 0.2 \times 0.3 + 0.3 \times 0.5 + 0.1 \times 0.8 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}$$

$$\begin{bmatrix} net_1^1 \\ net_2^1 \\ net_3^1 \end{bmatrix} = \underset{3\times 1}{\begin{bmatrix} 0.49 \\ 0.53 \\ 0.39 \end{bmatrix}} \xrightarrow{\text{ReLU}} \begin{bmatrix} O_1^1 \\ O_2^1 \\ O_3^1 \end{bmatrix} = \begin{bmatrix} 0.49 \\ 0.53 \\ 0.39 \end{bmatrix}$$

✔ ReLU(z) = max(0, z)

## Layer #2:

$$net_1^2 = \sum Wx + b = W_{11}^2 O_1^1 + W_{12}^2 O_2^1 + W_{13}^2 O_3^1 + b_1^2 = \left(\overset{0.049}{\cancel{0.1 \times 0.49}}\right) + \left(\overset{0.106}{\cancel{0.2 \times 0.53}}\right) + \left(\overset{0.078}{\cancel{0.2 \times 0.39}}\right) + 0.1$$

$$net_1^2 = 0.333 \xrightarrow{\tanh} O_1^2 = \tanh(net_1^2) = 0.321$$

✔ $\tanh(x) = \dfrac{\sinh(x)}{\cosh(x)} = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$

$$net_2^2 = \sum Wx + b = W_{21}^2 O_1^1 + W_{22}^2 O_2^1 + W_{23}^2 O_3^1 + b_2^2 = \left(\overset{0.196}{\cancel{0.4 \times 0.49}}\right) + \left(\overset{0.053}{\cancel{0.1 \times 0.53}}\right) + \left(\overset{0.078}{\cancel{0.2 \times 0.39}}\right) + 0.4$$

$$net_2^2 = 0.727 \xrightarrow{\tanh} O_2^2 = \tanh(net_2^2) = 0.621$$

Layer #3:

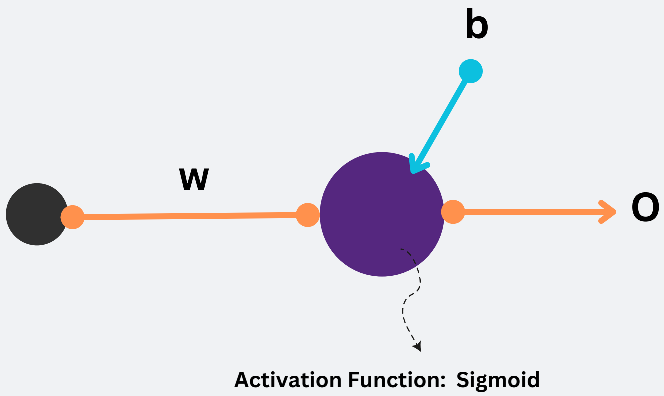
$$net_1^3 = \sum Wx + b = W_1^3 O_1^2 + W_2^3 O_2^2 + b^3 = (0.5 \times 0.321) + (0.1 \times 0.627) + 0.5 =$$

$$net_1^3 = 0.7232 \xrightarrow{\text{Sigmoid}} o^3 = \tanh(net_1^3) = 0.673$$

✓ $Sigmoid(x) = \sigma(x) = \dfrac{1}{1 + e^{-x}}$

## 2. Obtain new value of weight and bias after Gradient descent first iteration, use MSE as a loss function.

| x | y |
|---|---|
| 1.3 | 2.7 |
| 2.1 | 4.5 |
| 7.6 | 14.9 |



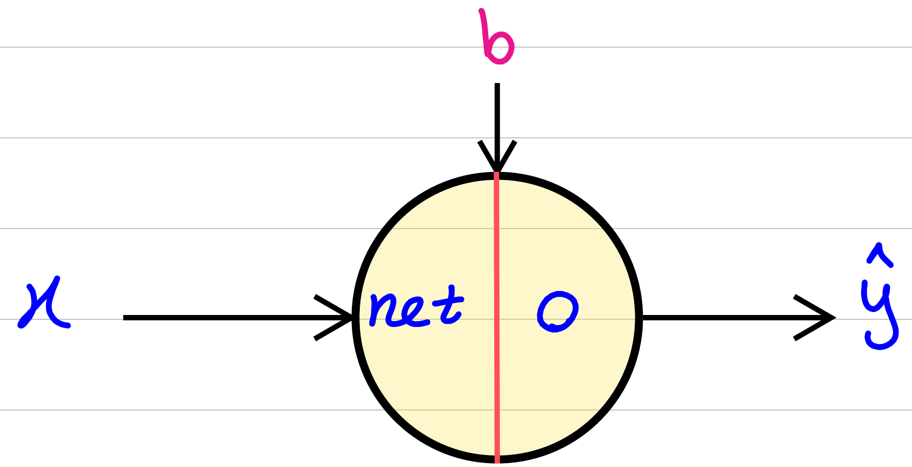

Initial weight value(weight have been initialized randomly) : 1.5

Initial bias value (weight have been initialized randomly) : 0.5

learning rate (alpha) = 0.1

400

# Gradient Descent in NN

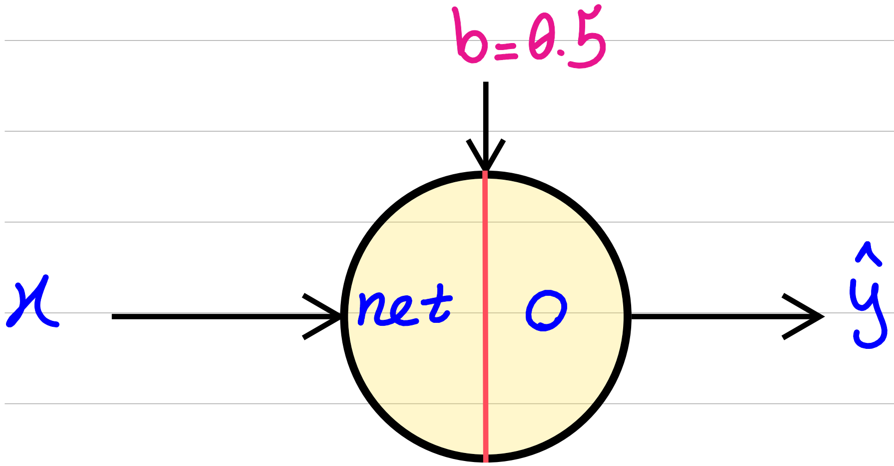



1. $\hat{y} = wx + b \longrightarrow$ (FF) خروجی مربوط به ورودی را محاسبه کنید

2. $Loss = \frac{1}{m}(\hat{y} - y)^2 \longrightarrow$ تابع هزینه را حساب کنید ، $\frac{\partial loss}{\partial w}$ , $\frac{\partial loss}{\partial b} = ?$

3. $w_{new} = w_{old} - \alpha \frac{\partial loss}{\partial w} \longrightarrow$ (BF) وزنها را آپدیت کنید

4. این مراحل را تا زمانیکه به حداقل Loss برسید تکرار کنید.




| $x$ | $y$ |
|---|---|
| 1.3 | 2.7 |
| 2.1 | 4.5 |
| 7.6 | 14.9 |

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

① $\hat{y} = Wx + b \rightsquigarrow net_1 = 1.5 \times 1.3 + 0.5 = 2.45 \xrightarrow{\text{Sigmoid}} \hat{y}_1 = 0.921$

$net_2 = 1.5 \times 2.1 + 0.5 = 3.65 \xrightarrow{\text{Sigmoid}} \hat{y}_2 = 0.975$

$net_3 = 1.5 \times 7.6 + 0.5 = 11.9 \xrightarrow{\text{Sigmoid}} \hat{y}_3 = 0.999$

② $$\text{Loss} = \text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 = \frac{1}{3}\left[(2.7-0.92)^2 + (4.5-0.97)^2 + (14.9-0.99)^2\right] = 69.7$$

$$\text{Loss} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 = \frac{1}{m}(Wx+b-y)^2 \longrightarrow \frac{\partial \text{loss}}{\partial W} = ?$$

$$\longrightarrow \frac{\partial \text{loss}}{\partial b} = ?$$

② $MSE = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2$

$$\frac{\partial MSE}{\partial w} = \underbrace{\frac{\partial MSE}{\partial \hat{y}}}_{\frac{2}{m}\Sigma(\hat{y}-y)} \underbrace{\frac{\partial \hat{y}}{\partial z}}_{\hat{y}(1-\hat{y})} \underbrace{\frac{\partial z}{\partial w}}_{x} = \frac{2}{m}\sum(\hat{y}-y)\,\hat{y}\,(1-\hat{y})\,x \qquad z_i = wx_i + b$$

$$\frac{\partial z}{\partial w} = (wx+b)' = x \qquad \frac{\partial \hat{y}}{\partial z} = \sigma(z)\,(1-\sigma(z)) \qquad * \hat{y} = \sigma(z_i) = \frac{1}{1+e^{-z_i}}$$

$$* \longrightarrow \frac{\partial \hat{y}}{\partial z} = \hat{y}\,(1-\hat{y})$$

$$\frac{\partial MSE}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}\left(\frac{1}{m}\sum(\hat{y}-y)^2\right) = \frac{2}{m}\sum(\hat{y}-y)(1)$$

③ $w_{new} = w_{old} - \alpha \frac{\partial loss}{\partial w} =$

$$\frac{2}{m}\sum(\hat{y}-y)\,\hat{y}\,(1-\hat{y})\,x \longrightarrow \begin{cases} 2\overset{-3.558}{\cancel{(0.921-2.7)}} \times \cancel{0.921} \times \overset{0.073}{\cancel{(1-0.921)}} \times 1.3 = -0.337 \\ 2\overset{-7.05}{\cancel{(0.975-4.5)}} \times \cancel{0.975} \times \overset{0.024}{\cancel{(1-0.975)}} \times 2.1 = -0.356 \\ 2\overset{-27.8}{\cancel{(0.999-14.9)}} \times \cancel{0.999} \times \overset{0.001}{\cancel{(1-0.999)}} \times 7.6 = -0.211 \end{cases} \xrightarrow{AVG} \frac{1}{3}\overset{-0.904}{\left[(-0.337)+(-0.356)+(-0.211)\right]} = -0.301$$

$$w_{new} = w_{old} - \alpha\frac{\partial l}{\partial w} \Rightarrow w_{new} = 1.5 - (0.1 \times (-0.301)) = 1.5301$$

② $MSE = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2$

$$\frac{\partial MSE}{\partial b} = \frac{\partial MSE}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z} \frac{\partial z}{\partial b} = \frac{2}{m}(\hat{y}-y)\,\hat{y}\,(1-\hat{y})$$

$$\frac{\partial MSE}{\partial \hat{y}} = \frac{2}{m}(\hat{y}-y)(1) \qquad \frac{\partial \hat{y}}{\partial z} = \hat{y}(1-\hat{y}) \qquad \frac{\partial z}{\partial b} = 1$$

③ $b_{new} = b_{old} - \alpha \frac{\partial loss}{\partial b} =$

$\frac{2}{m}\sum(\hat{y}-y)\,\hat{y}\,(1-\hat{y})$

$2(0.921-2.7)\times 0.921\times(1-0.921) = -0.258$

$2(0.975-4.5)\times 0.975\times(1-0.975) = -0.172$

$2(0.999-14.5)\times 0.999\times(1-0.999) = -0.026$

$\xrightarrow{AVG} \frac{1}{3}\times(-0.456) = -0.152$

$$b_{new} = b_{old} - \alpha\frac{\partial L}{\partial b} \Rightarrow b_{new} = 0.5 - (0.1\times(-0.152)) = 0.5152$$