Forward Propagation - **Hidden layer - 1**

$O_1 = x_1(w_{11}) + x_2(w_{16}) + x_3(w_{111}) + \text{Bias}$
$= 1000(1) + 1(1) + 1(1) + 0.01$
$= 1002.01$

$O_2 = x_1(w_{12}) + x_2(w_{17}) + x_3(w_{112}) + \text{Bias}.$
$= 1000(2) + 1(2) + 1(2) + 0.01$
$= 2004.01$

$O_3 = x_1(w_{13}) + x_2(w_{18}) + x_3(w_{113}) + \text{Bias}$
$= 1000(3) + 1(3) + 1(3) + 0.01$
$= 3006.01$

$O_4 = x_1(w_{14}) + x_2(w_{19}) + x_3(w_{114}) + \text{Bias}.$
$= 1000(4) + 1(4) + 1(4) + 0.01 = 4008.01$

$O_5 = x_1(w_{15}) + x_2(w_{20}) + x_3(w_{115}) + \text{Bias}$
$= 1000(5) + 1(5) + 1(5) + 0.01 = 1010.01$

Assuming
$w_{11} = 1$, $w_{16} = 1$
$w_{12} = 2$ $w_{17} = 2$
$w_{13} = 3$ $w_{18} = 3$
$w_{14} = 4$ $w_{19} = 4$
$w_{15} = 5$ $w_{20} = 5$

$w_{111} = 1$
$w_{112} = 2$
$w_{113} = 3$
$w_{114} = 4$
$w_{115} = 5$
$x_1 = 1000$
$x_2 = 1$
$x_3 = 1$
Bias = 0.01

For hidden layer will use sigmoid activation fun.

Hidden layer - 2

$O_6 = O_1(w_{21}) + O_2(w_{26}) + O_3(w_{31}) + O_4(w_{36}) + O_5(w_{41}) + \text{Bias}$
$= 1002.01(1) + 2004.01(1) + 3006.01(1) +$
$+ 4008.01(1) + 1010.01(1)$
$=$

$O_7 = O_1(w_{22}) + O_2(w_{27}) + O_3(w_{32}) + O_4(w_{37})$
$+ O_5(w_{42}) + \text{Bias}.$
$= 1002.01(2) + 2004.01(2) + 3006.01(2) +$
$+ 4008.01(2) + 1010.01(2) + 0.01$

$O_8 = O_1(w_{23}) + O_2(w_{28}) + O_3(w_{33}) + O_4(w_{38})$
$+ O_5(w_{43}) + \text{Bais}$

$O_9 = O_1(w_{24}) + O_2(w_{29}) + O_3(w_{34}) + O_4(w_{39})$
$+ O_5(w_{44}) + \text{Bias}$

$O_{10} = O_1(w_{25}) + O_2(w_{30}) + O_3(w_{35}) + O_4(w_{40}) + O_5(w_{45}) + \text{Bais}$

Assuming
$w_{21} = 1$, $w_{26} = 1$
$w_{22} = 2$ $w_{27} = 2$
$w_{23} = 3$ $w_{28} = 3$
$w_{24} = 4$ $w_{29} = 4$
$w_{25} = 5$ $w_{30} = 5$

$w_{31} = 1$ $w_{36} = 1$
$w_{32} = 2$ $w_{37} = 2$
$w_{33} = 3$ $w_{38} = 3$
$w_{34} = 4$ $w_{39} = 4$
$w_{35} = 5$ $w_{40} = 5$

$w_{41} = 1$
$w_{42} = 2$
$w_{43} = 3$
$w_{44} = 4$
$w_{45} = 5$

$$O_{11} = O_6(W_{50}) + O_7(W_{51}) + O_8(W_{52}) + O_9(W_{53}) + O_{10}(W_{54}) + \text{Bias}$$

$$= O_6(1) + O_7(2) + O_8(3) + O_9(4) + O_{10}(5) + 0.01$$

$$= 100(1) + 200(2) + 300(3) + 400(4) + 500(5) + 0.01$$

$$= 100 + 400 + 900 + 1600 + 2500 + 0.01$$

$$= 5500 + 0.01$$

$$= 5500.01$$

Assuming
$W_{50} = 1$
$W_{51} = 2$
$W_{52} = 3$
$W_{53} = 4$
$W_{55} = 5$
Bias $= 0.01$

Assume
$O_6 = 100$
$O_7 = 200$
$O_8 = 300$
$O_9 = 400$
$O_{10} = 50$

using Relu activation $=$ Relu(5500.01)

$O_{11} \approx 65$ L.

Actual $= 50$L.

Prediction $\doteq 65$ L

Loss $= 65 - 50 = 15$ L

After Calculation hidden layer

Back Propagation - chain Rule - (output layer to 2nd hidden layer)

$$W_{50} = \frac{dL}{dW_{50}} = \frac{dL}{dO_{11}} * \frac{dO_{11}}{dW_{50}}$$

$$W_{51} = \frac{dL}{dW_{51}} = \frac{dL}{dO_{11}} * \frac{dO_{11}}{dW_{51}}$$

$$W_{52} = \frac{dL}{dW_{52}} = \frac{dL}{dO_{11}} * \frac{dO_{11}}{dW_{52}}$$

$$W_{53} = \frac{dL}{dW_{53}} = \frac{dL}{dO_{11}} * \frac{dO_{11}}{dW_{53}}$$

$$W_{54} = \frac{dL}{dW_{54}} = \frac{dL}{dO_{11}} * \frac{dO_{11}}{dW_{53}}$$

LR = Learning Rate
Range (0.001 - 1)
= 0.001.

Update of weight -

$$W_{50}(new) = W_{50}(old) - LR\left(\frac{dL}{dW_{50}}\right)$$

$$W_{51}(new) = W_{51}(old) - LR\left(\frac{dL}{dW_{51}}\right)$$

$$W_{52}(new) = W_{52}(old) - LR\left(\frac{dL}{dW_{52}}\right)$$

$$W_{53}(new) = W_{53}(old) - LR\left(\frac{dL}{dW_{53}}\right)$$

$$W_{54}(new) = W_{54}(old) - LR\left(\frac{dL}{dW_{54}}\right)$$

Back Propagation - chain Rule (2nd hidden layer to 1st hidden layer)

$$W_{21} = \frac{dL}{dW_{21}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_6} \times \frac{dO_6}{dW_{21}}$$

$$W_{22} = \frac{dL}{dW_{22}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_7} \times \frac{dO_7}{dO_{22}}$$

$$W_{23} = \frac{dL}{dW_{23}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_8} \times \frac{dO_8}{dW_{23}}$$

$$W_{24} = \frac{dL}{dW_{24}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_9} \times \frac{dO_9}{dW_{24}}$$

$$W_{25} = \frac{dL}{dW_{25}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_{10}} \times \frac{dO_{10}}{dW_{25}}$$

$$W_{26} = \frac{dL}{dW_{26}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_6} \times \frac{dO_6}{dW_{26}}$$

$$W_{45} = \frac{dL}{dW_{45}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_{10}} \times \frac{dO_{11}}{dW_{45}}$$

Update of weight

$$W_{21}(new) = W_{21}(old) - LR\left(\frac{dL}{dW_{21}}\right)$$

$$W_{22}(new) = W_{22}(old) - LR\left(\frac{dL}{dW_{22}}\right)$$

LR = Learning Rate
Range → 0.001 - 1
= 0.001

$$W_{45}(new) = W_{45}(old) - LR\left(\frac{dL}{dW_{45}}\right)$$

$$W_{11} = \frac{dL}{dW_{11}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_6} \times \frac{dO_6}{dO_1} \times \frac{dO_1}{dW_{11}}$$

$$W_{16} = \frac{dL}{dW_{16}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_7} \times \frac{dO_7}{dO_1} \times \frac{dO_1}{dW_{16}}$$

$$\vdots$$

$$W_{115} = \frac{dL}{dW_{115}} = \frac{dL}{dO_{11}} \times \frac{dO_{11}}{dO_{10}} \times \frac{dO_{10}}{dO_5} \times \frac{dO_5}{dW_{115}}$$

update of weight -

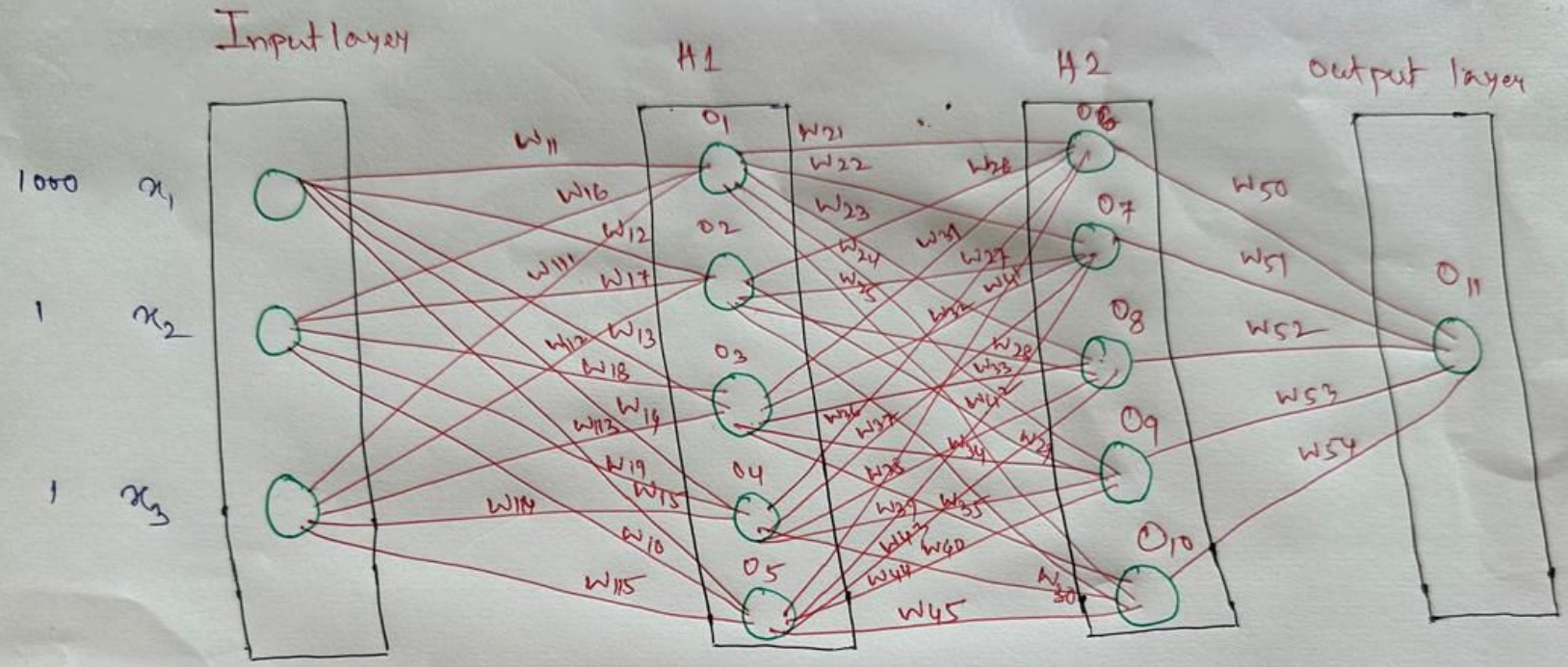
$$W_{11}(new) = W_{11}(old) - LR\left(\frac{dL}{dW_{11}}\right)$$

$$W_{12}(new) = W_{12}(old) - LR\left(\frac{dL}{dW_{12}}\right)$$

$$\vdots$$

$$W_{115}(new) = W_{115}(old) - LR\left(\frac{dL}{dW_{115}}\right)$$

Consider a 1 epoch (1 forward Propagation & 1 backward Propagation)



Use Case - House value Prediction.

1 input layer, 2 hidden layer, 1 output layer.

Input contain 3 feature x_1 = 1000 sq.ft (area), x_2 = 1 (locality), x_3 = 1 quality of raw material.

Each hidden layer contain 5 NN.

Since its a regression Problem output is 1 (Buy or not).

Actual Price = 50 Lacks