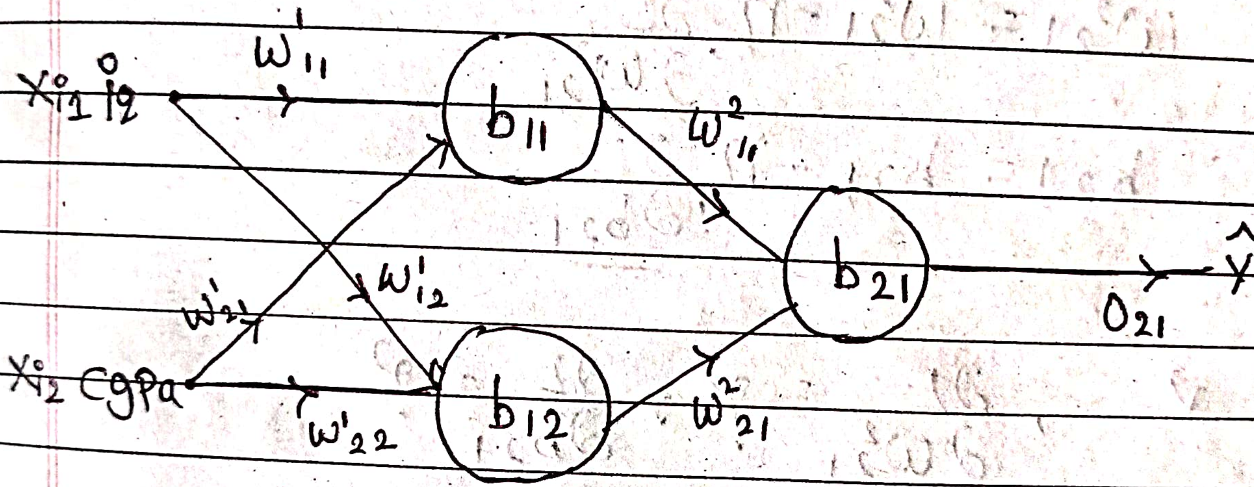
Back Propagation
ANN

Linear Regression Problem

Dataset

iq	cgpa	package
80	8	8
70	7	7
60	6	6
50	5	5



training

Step 1: Initialize w's & b (Weights & biases), $w=1, b=0$

Step 2: Predict output ($\hat{y}$) (forward propagation)

Step 3: calculate loss function (mse)

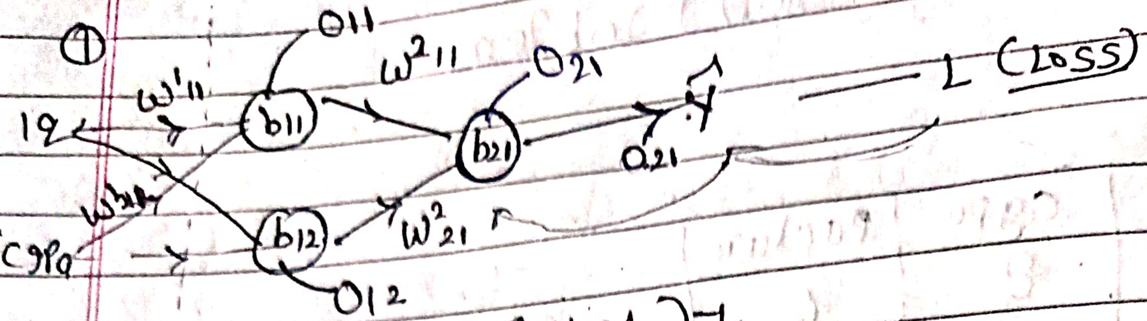
Step 4: Update weights & biases using gradient descent,

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

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

break NN :

①



Update formula (W & b) ↓

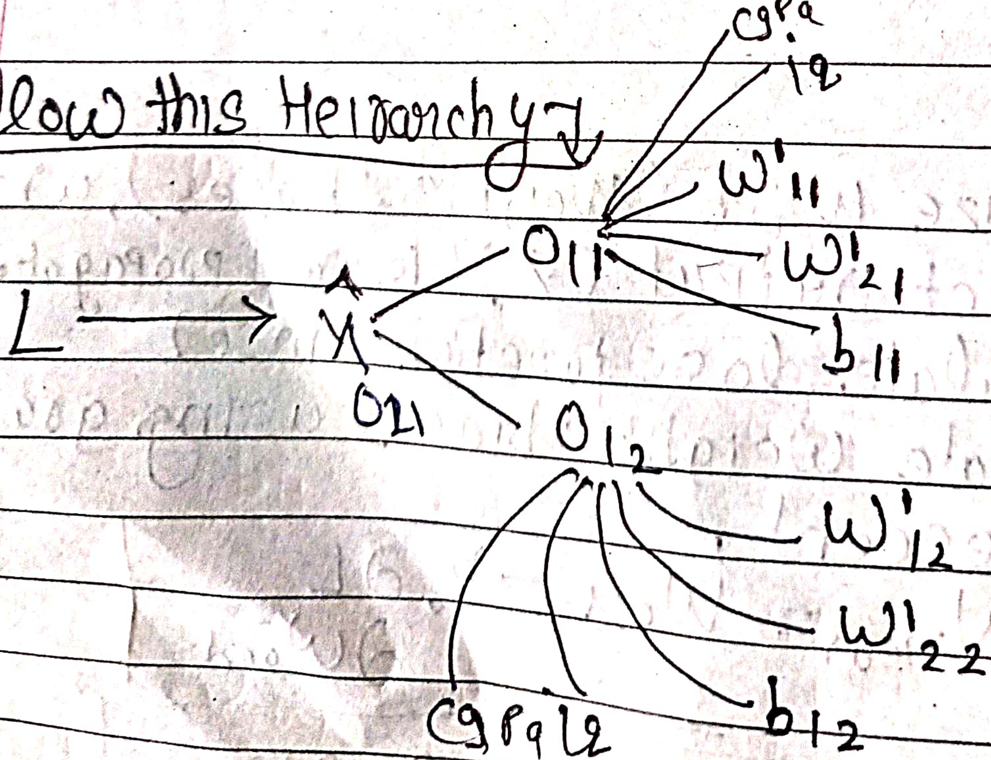
$$w^2_{11} = w^2_{11} - \eta \frac{\partial L}{\partial w^2_{11}}$$

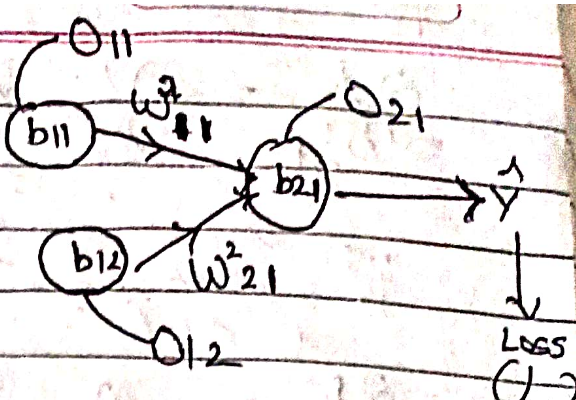
$$w^2_{21} = w^2_{21} - \eta \frac{\partial L}{\partial w^2_{21}}$$

$$b_{21} = b_{21} - \eta \frac{\partial L}{\partial b_{21}}$$

$\frac{\partial L}{\partial w^2_{11}} = ?$, $\frac{\partial L}{\partial w^2_{21}} = ?$, $\frac{\partial L}{\partial b_{21}} = ?$

Follow this Heirarchy ↓





$$\frac{\partial L}{\partial w^2_{11}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{w^2_{11}}$$

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

$$= -2(Y - \hat{y})$$

$$\frac{\partial \hat{y}}{\partial w^2_{11}} = \frac{\partial}{\partial w^2_{11}} \left[w^2_{11} O_{11} + w^2_{21} O_{12} + b_{21} \right]$$

$$= O_{11}$$

$$\frac{\partial L}{\partial w^2_{11}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial w^2_{11}}$$

$$= -2(Y - \hat{y}) O_{11}$$

$$\boxed{\frac{\partial L}{\partial w^2_{11}} = -2(Y - \hat{y}) O_{11}}$$

$$\frac{\partial L}{\partial w^2_{21}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial w^2_{21}}$$

$$\frac{\partial \hat{y}}{\partial w^2_{21}} = \frac{\partial}{\partial w^2_{21}} \left[w^2_{11} O_{11} + w^2_{21} O_{12} + b_{21} \right]$$

$$= O_{12}$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} = -2(Y + \hat{y})}$$

$$\boxed{\frac{\partial L}{\partial w^2_{21}} = -2(Y - \hat{y}) O_{12}}$$

Date: / /

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

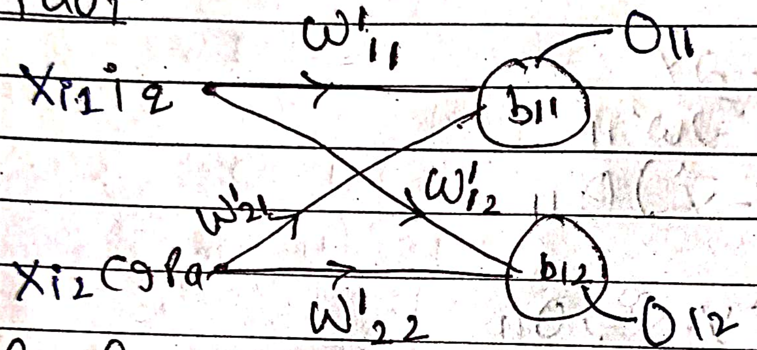
$$\because \frac{\partial L}{\partial \hat{y}} = -2(y-\hat{y})$$

$$\frac{\partial \hat{y}}{\partial b_{21}} = \frac{\partial}{\partial b_{21}}\left[w^2_{11} O_{11} + w^2_{21} O_{12} + b_{21}\right]$$

$$= 1$$

$$\boxed{\frac{\partial L}{\partial b_{21}} = -2(y-\hat{y})}$$

IInd part



formula for updation,

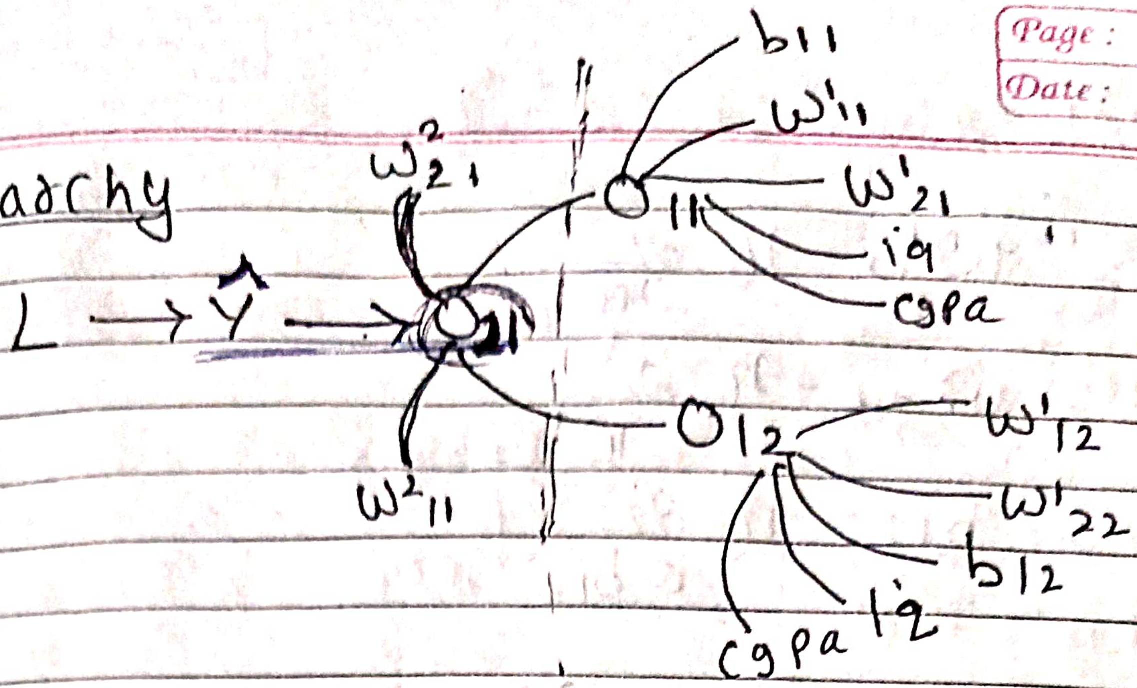
$$w'_{11} = w'_{11} - \eta \frac{\partial L}{\partial w'_{11}}$$

$$w'_{21} = w'_{21} - \eta \frac{\partial L}{\partial w'_{21}}$$

$$b_{11} = b_{11} - \eta \frac{\partial L}{\partial b_{11}}$$

$$\frac{\partial L}{\partial w'_{11}} = ? \quad , \quad \frac{\partial L}{\partial w'_{21}} = ? \quad \frac{\partial L}{\partial b_{11}} = ?$$

Heioarchy



$$\frac{\partial L}{\partial w'_{11}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial w'_{11}}$$

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

$$\frac{\partial \hat{y}}{\partial O_{11}} = \frac{\partial}{\partial O_{11}}\left[w^2_{11} O_{11} + w^2_{21} O_{12} + b_{21}\right]$$

$$= w^2_{11}$$

$$\frac{\partial O_{11}}{\partial w'_{11}} = \frac{\partial}{\partial w'_{11}}\left[X_{i1} w'_{11} + X_{i2} w'_{21} + b_{11}\right]$$

$$= X_{i1}$$

$$\boxed{\frac{\partial O_{11}}{\partial w'_{11}} = -2(y-\hat{y}) w^2_{11} X_{i1}}$$

Similarly

$$\boxed{\frac{\partial L}{\partial w'_{21}} = -2(y-\hat{y}) w^2_{11} X_{i2}}$$

$$\boxed{\frac{\partial L}{\partial b_{11}} = -2(y-\hat{y}) w^2_{11}}$$

$$w'_{12} = w'_{12} - \eta \frac{\partial L}{\partial w'_{12}}$$

$$w'_{22} = w'_{22} - \eta \frac{\partial L}{\partial w'_{22}}$$

$$b_{12} = b_{12} - \eta \frac{\partial L}{\partial b_{12}}$$

$$\frac{\partial L}{\partial w'_{12}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial w'_{12}}$$

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

$$\frac{\partial \hat{y}}{\partial O_{12}} = \frac{\partial}{\partial O_{12}} \left[w^2_{11} O_{11} + w^2_{21} O_{12} + b_{21} \right]$$

$$= w^2_{21}$$

$$\frac{\partial O_{12}}{\partial w'_{12}} = \frac{\partial}{\partial w'_{12}} \left[X_{i1} w'_{12} + X_{i2} w'_{22} + b_{12} \right]$$

$$= \cancel{X_{i1}}\ X_{i1}$$

$$\boxed{\frac{\partial L}{\partial w'_{12}} = -2(y-\hat{y}) w^2_{21} X_{i1}}$$

similarly,

$$\boxed{\frac{\partial L}{\partial w'_{22}} = -2(y-\hat{y}) w^2_{21} X_{i2}}$$

$$\boxed{\frac{\partial L}{\partial b_{12}} = -2(y-\hat{y}) w^2_{21}}$$ Ans.

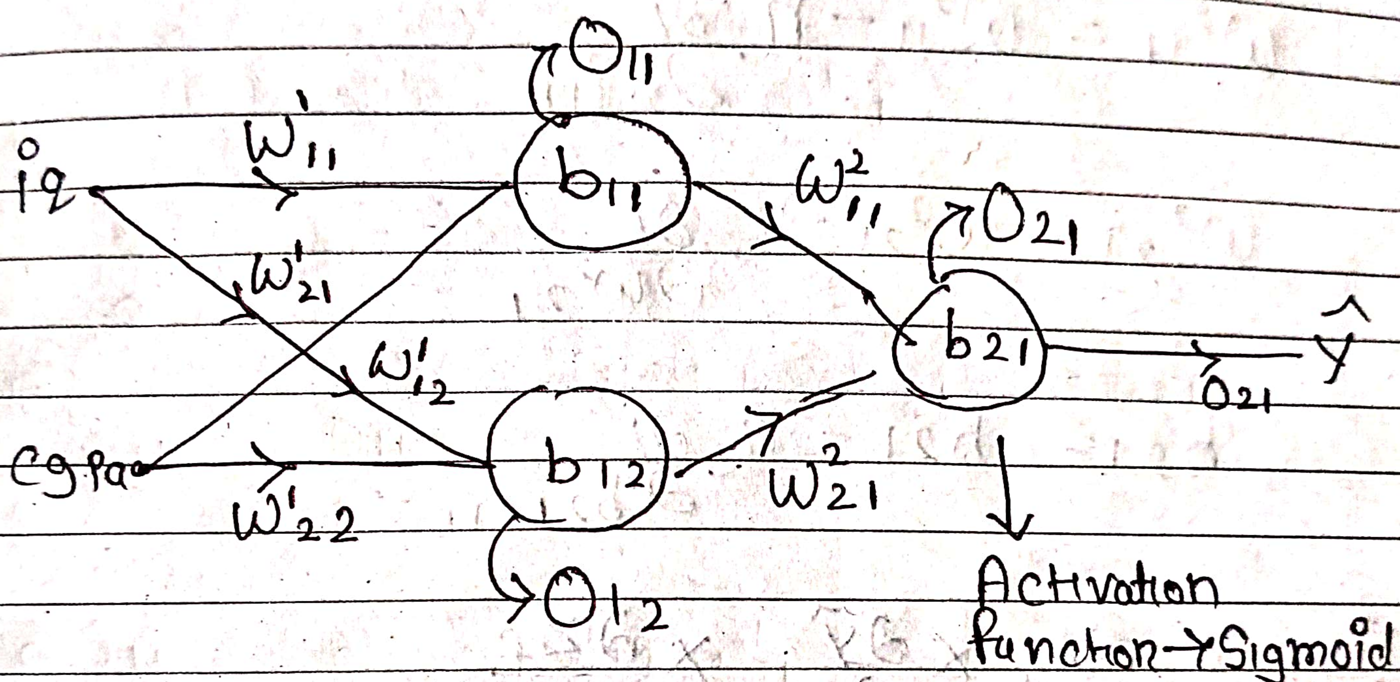
Date: / /

Classification problem

Dataset

1: yes, 0: No

iq	cgpa	placed
80	8	1
70	7	1
60	6	0
50	5	0



loss function [binary cross entropy]

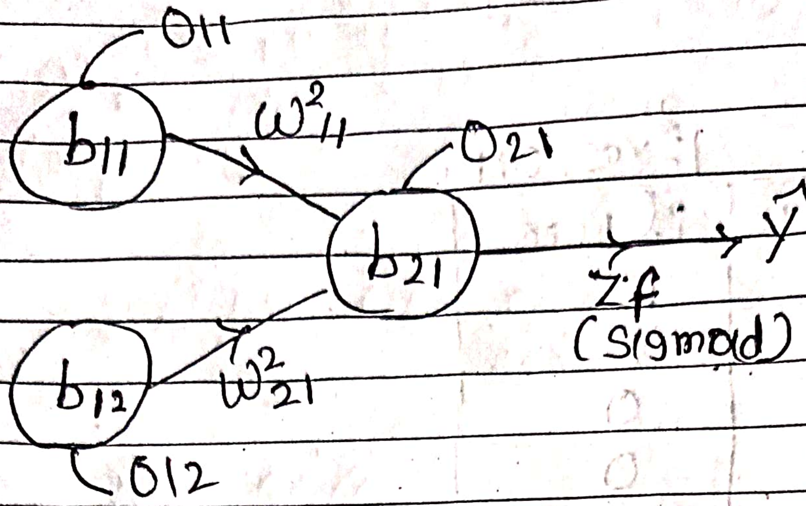
$$\boxed{L = -y\log\hat{y} - (1-y)\log(1-\hat{y})}$$

where, y → actual label

$\hat{y}$ → predicted label

Date: / /

Split N.N (Neural Network)



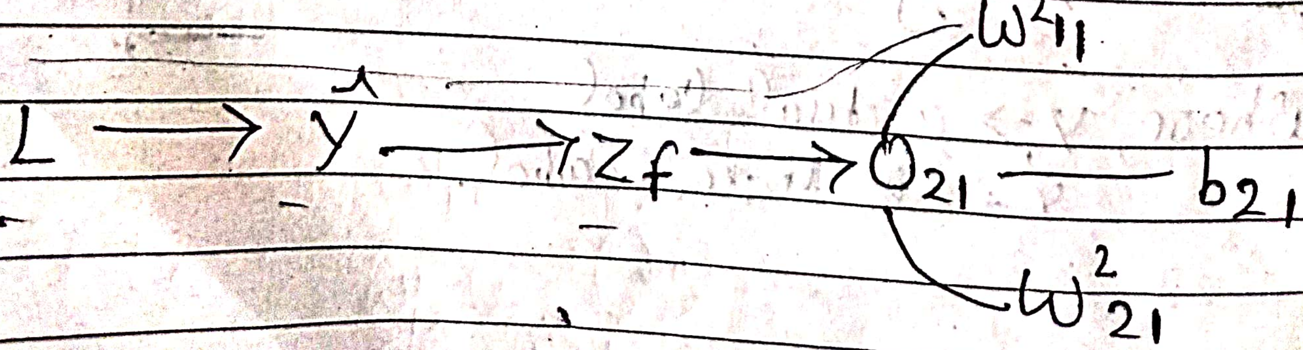
$$w^2_{11} = w^2_{11} - \eta \frac{\partial L}{\partial w^2_{11}}$$

$$w^2_{21} = w^2_{21} - \eta \frac{\partial L}{\partial w^2_{21}}$$

$$b_{21} = b_{21} - \eta \frac{\partial L}{\partial b_{21}}$$

$$\frac{\partial L}{\partial w^2_{11}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial Z_f} \times \frac{\partial Z_f}{\partial w^2_{11}}$$

Hierarchy follows,



Date: / /

$$\frac{\partial L}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}\left(-y\log(\hat{y}) - (1-y)\log(1-\hat{y})\right)$$

$$= \cancel{\frac{\partial}{\partial \hat{y}}((1-\hat{y}}$$

$$= -y\frac{\partial}{\partial \hat{y}}(\log(\hat{y})) - (1-y)\frac{\partial}{\partial \hat{y}}(\log(1-\hat{y}))$$

$$= -\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\frac{\partial}{\partial \hat{y}}(1-\hat{y})$$

$$= -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}$$

$$= \frac{-y(1-\hat{y}) + (1-y)\hat{y}}{\hat{y}(1-\hat{y})}$$

$$= \frac{-y + \cancel{y\hat{y}} + \hat{y} - \cancel{y\hat{y}}}{\hat{y}(1-\hat{y})}$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} = -\frac{(y-\hat{y})}{\hat{y}(1-\hat{y})}}$$

$$\frac{\partial \hat{y}}{\partial z_f} = \frac{\partial}{\partial z_f}(\hat{y}) \qquad \because \hat{y} = \sigma(z)$$

Derivation of

$$= \frac{\partial}{\partial z_f}(\sigma(z)) \qquad \sigma(z) = \sigma(z)[1-\sigma(z)]$$

$$= \sigma(z)[1-\sigma(z)]$$

$$\because \sigma(z) = \hat{y}$$

$$\boxed{\frac{\partial \hat{y}}{\partial z_f} \Rightarrow \hat{y}(1-\hat{y})}$$

$$\because z_f = O_{11}W^2_{11} + O_{12}W^2_{21} + b_{21}$$

$$\frac{\partial z_f}{\partial W^2_{11}} \Rightarrow \frac{\partial}{\partial W^2_{11}}(z_f).$$

$$\frac{\partial}{\partial w^2_{11}}\left(O_{11}\underline{w^2_{11}} + O_{12}w^2_{21} + b_{21}\right)$$

$$\Rightarrow \underline{O_{11}}$$

$$\frac{\partial L}{\partial w^2_{11}} \Rightarrow \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_f} \times \frac{\partial z_f}{\partial w^2_{11}}$$

$$\Rightarrow -\cancel{\frac{(y-\hat{y})}{\hat{y}(1-\hat{y})}} \times \cancel{\hat{y}(1-\hat{y})} \times O_{11}$$

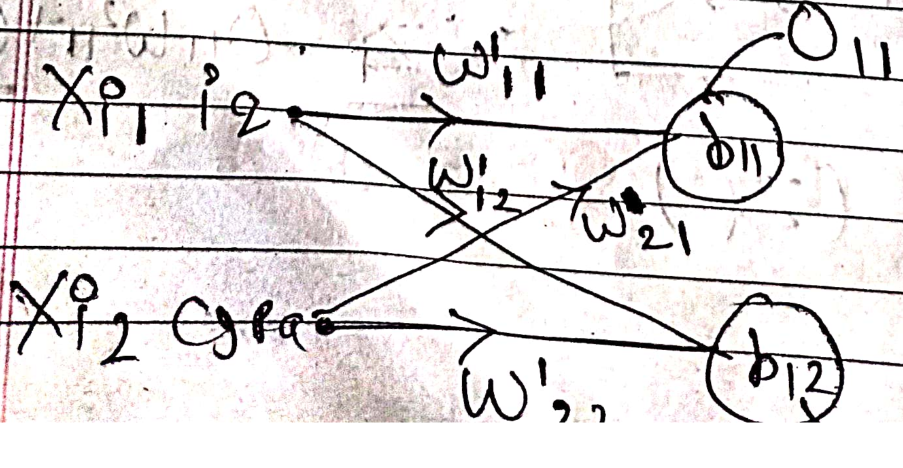
$$\boxed{\frac{\partial L}{\partial w^2_{11}} \Rightarrow -(y-\hat{y})\,O_{11}}$$

<u>Similarly</u>

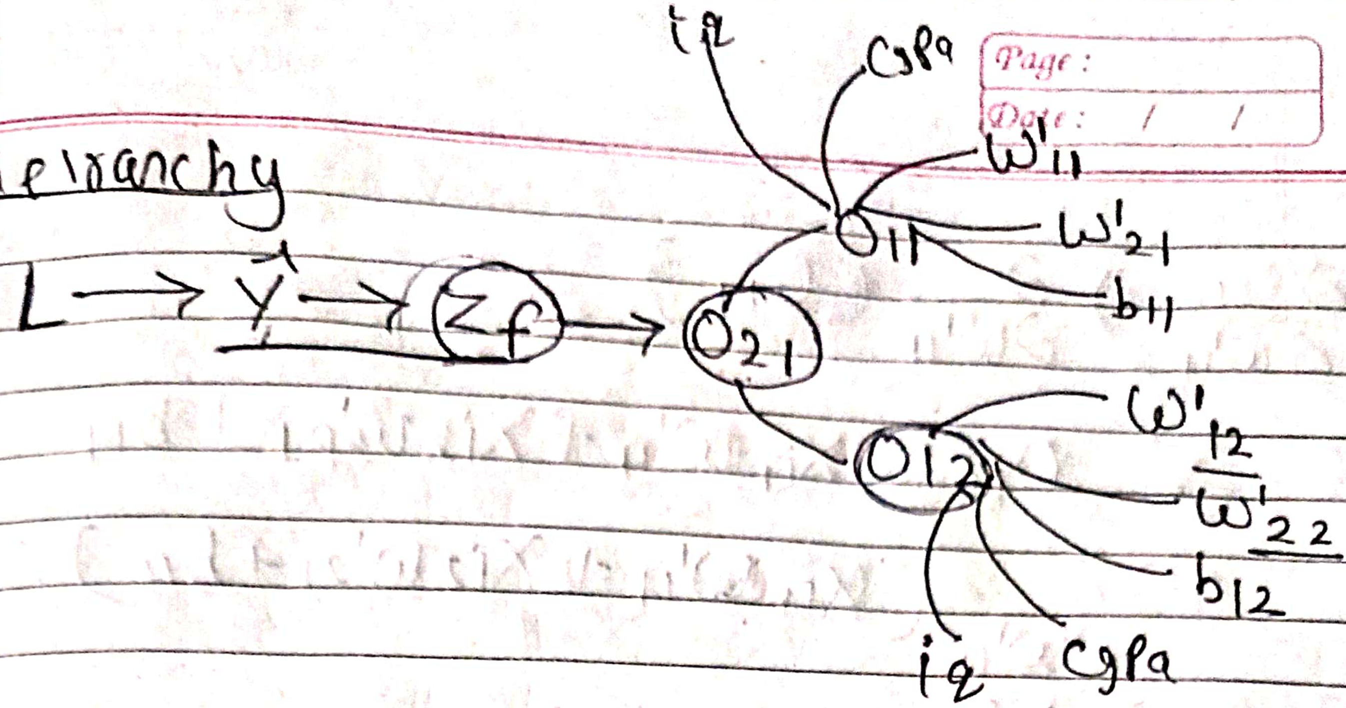
$$\boxed{\frac{\partial L}{\partial w^2_{21}} = -(y-\hat{y})\,O_{12}}$$

$$\boxed{\frac{\partial L}{\partial b_{21}} = -(y-\hat{y})}$$

<u>IInd part</u>



Heirarchy



$$w'_{11} = w'_{11} - \eta \frac{\partial L}{\partial w'_{11}}$$

$$w'_{21} = w'_{21} - \eta \frac{\partial L}{\partial w'_{21}}$$

$$b_{11} = b_{11} - \eta \frac{\partial L}{\partial b_{21}}$$

$$\frac{\partial L}{\partial w'_{11}} = \underbrace{\frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial Zf}} \times \frac{\partial Zf}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial w'_{11}}$$

$$\therefore \boxed{\begin{aligned} \frac{\partial L}{\partial \hat{y}} &= -\frac{(y - \hat{y})}{\hat{y}(1-\hat{y})} \\ \frac{\partial \hat{y}}{\partial Zf} &= \hat{y}(1-\hat{y}) \end{aligned}}$$

$$\frac{\partial Zf}{\partial O_{11}} = \frac{\partial}{\partial O_{11}} (Zf)$$

$$= \frac{\partial}{\partial O_{11}} (O_{11} w^2_{11} + O_{12} w^2_{21} + b_{21})$$

$$\Rightarrow w^2_{11}$$

$$\frac{\partial O_{11}}{\partial w'_{11}} \Rightarrow \frac{\partial}{\partial w'_{11}} (O_{11})$$

$$\therefore O_{11} = Xi_1 w'_{11} + Xi_2 w'_{21} + b_{11}$$

$$\Rightarrow \frac{\partial}{\partial w'_{11}} (Xi_1 w'_{11} + Xi_2 w'_{21} + b_{11})$$

$$\Rightarrow Xi_1$$

$$\frac{\partial L}{\partial w'_{11}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial Z_f} \times \frac{\partial Z_f}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial w'_{11}}$$

$$\Rightarrow -\frac{(y-\hat{y})}{\cancel{\hat{y}(1-\hat{y})}} \times \cancel{(\hat{y})(1-\hat{y})} \times w^2_{11} \times Xi_1$$

$$\boxed{\frac{\partial L}{\partial w'_{11}} \Rightarrow -(y-\hat{y})\, w^2_{11}\, Xi_1}$$

Similarly

$$\boxed{\frac{\partial L}{\partial w'_{21}} = -(y-\hat{y})\, w^2_{11}\, Xi_2}$$

$$\boxed{\frac{\partial L}{\partial b_{11}} = -(y-\hat{y})\, w^2_{11}}$$

Last

$$w'_{12} = w'_{12} - \eta \frac{\partial L}{\partial w'_{12}}$$

Date: / /

$$w'_{22} = w'_{22} - \eta \frac{\partial L}{\partial w'_{22}}$$

$$b_{12} = b_{12} - \eta \frac{\partial L}{\partial b_{12}}$$

$$\frac{\partial L}{\partial w'_{12}} = \underbrace{\frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_f}} \times \frac{\partial z_f}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial w'_{12}}$$

$$\therefore \boxed{\begin{array}{l} \dfrac{\partial L}{\partial \hat{y}} = -\dfrac{(y-\hat{y})}{\hat{y}(1-\hat{y})} \\ \dfrac{\partial \hat{y}}{\partial z_f} = \hat{y}(1-\hat{y}) \end{array}}$$

$$\frac{\partial z_f}{\partial O_{12}} = \frac{\partial}{\partial O_{12}}[z_f]$$

$$\Rightarrow \frac{\partial}{\partial O_{12}}\left[O_{11}w^2_{11} + O_{12}w^2_{21} + b_{21}\right]$$

$$\Rightarrow w^2_{21}$$

$$\frac{\partial O_{12}}{\partial w'_{12}} \Rightarrow \frac{\partial}{\partial w'_{12}}(O_{12})$$

$$\therefore O_{12} = X_{i1}w'_{12} + X_{i2}w'_{22} + b_{12}$$

$$\Rightarrow \frac{\partial}{\partial w'_{12}}(X_{i1}w'_{12} + X_{i2}w'_{22} + b_{12})$$

$$\Rightarrow X_{i1}$$

$$\frac{\partial L}{\partial w'_{12}} \Rightarrow \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial Zf} \times \frac{\partial Zf}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial w'_{12}}$$

$$\Rightarrow -\frac{(y-\hat{y})}{\cancel{\hat{y}(1-\hat{y})}} \times \cancel{\hat{y}(1-\hat{y})} \times w^2_{21} \times x_{i1}$$

$$\boxed{\frac{\partial L}{\partial w'_{12}} \Rightarrow -(y-\hat{y})\, w^2_{21}\, x_{i1}}$$

Similarly

$$\boxed{\frac{\partial L}{\partial w'_{22}} = -(y-\hat{y})\, w^2_{21}\, x_{i2}}$$

$$\boxed{\frac{\partial L}{\partial b_{12}} = -(y-\hat{y})\, w^2_{21}}$$