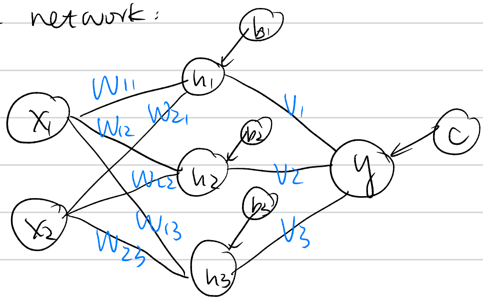
Q1. Feed Forward

1. Draw a network:



2. Write out mathematical equation for the output

$$h_1 = \max(0, W_{11}X_1 + W_{21}X_2 + b_1)$$
$$h_2 = \max(0, W_{12}X_1 + W_{22}X_2 + b_2)$$
$$h_3 = \max(0, W_{13}X_1 + W_{23}X_2 + b_3)$$
$$y = \text{sigmoid}(V_1h_1 + V_2h_2 + V_3h_3 + C)$$
$$= \frac{1}{1+e^{-(V_1h_1+V_2h_2+V_3h_3+C)}}$$

Q2. Gradient Descent

1. $$\frac{\partial f(x,y)}{\partial x} = \frac{(1-x^3)+100(y^2-x)^2}{\partial x} = -3x^2 - 200(y^2-x)$$

$$\frac{\partial f(x,y)}{\partial y} = \frac{(1-x^3)+100(y^2-x)^2}{\partial y} = 200(y^2-x)\cdot 2y = 400y(y^2-x)$$

Q3. Backprop.

1. Derive expressions of the gradient of the Loss function

$$\text{Loss} = -(y\log(\hat{y}) + (1-y)\log(1-\hat{y}))$$

$$= -y\log\left(\frac{1}{1+e^{-(V_1h_1+V_2h_2+V_3h_3+C)}}\right) - ((1-y)(1-\log\frac{1}{1+e^{-(V_1h_1+V_2h_2+V_3h_3+C)}}))$$

$$\text{Sigmoid}: \frac{1}{1+e^{-z}} \qquad z = v_1h_1 + v_2h_2 + v_3h_3 + c \qquad in_i = \sum_{j=1,2} W_{ij}x_i + b_i$$

$$\frac{\partial L}{\partial v_1} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial v_1} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))(-h_1)$$

$$\frac{\partial L}{\partial v_2} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))(-h_2)$$

$$\frac{\partial L}{\partial v_3} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))(-h_3)$$

$$\frac{\partial L}{\partial W_{11}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_1}\frac{\partial out_1}{\partial in_1}\frac{\partial in_1}{\partial W_{11}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_1v_1, & in_1 > 0 \\ 0, & in_1 < 0 \end{cases}$$

$$\frac{\partial L}{\partial W_{12}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_2}\frac{\partial out_2}{\partial in_2}\frac{\partial in_2}{\partial W_{12}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_1v_2, & in_2 > 0 \\ 0, & in_2 < 0 \end{cases}$$

$$\frac{\partial L}{\partial W_{13}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_3}\frac{\partial out_3}{\partial in_3}\frac{\partial in_3}{\partial W_{13}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_1v_3, & in_3 > 0 \\ 0, & in_3 < 0 \end{cases}$$

$$\frac{\partial L}{\partial W_{21}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_1}\frac{\partial out_1}{\partial in_1}\frac{\partial in_1}{\partial W_{21}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_2v_1, & in_2 > 0 \\ 0, & in_2 < 0 \end{cases}$$

$$\frac{\partial L}{\partial W_{22}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_2}\frac{\partial out_2}{\partial in_2}\frac{\partial in_2}{\partial W_{22}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_2v_2, & in_2 > 0 \\ 0, & in_2 < 0 \end{cases}$$

$$\frac{\partial L}{\partial W_{23}} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_3}\frac{\partial out_3}{\partial in_3}\frac{\partial in_3}{\partial W_{23}} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z))x_2v_3, & in_2 > 0 \\ 0, & in_2 < 0 \end{cases}$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_1}\frac{\partial out_1}{\partial in_1}\frac{\partial in_1}{\partial b_1} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z)), & in_1 > 0 \\ 0, & in_1 < 0 \end{cases}$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial out_2}\frac{\partial out_2}{\partial in_2}\frac{\partial in_2}{\partial b_2} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z)), & in_2 > 0 \\ 0, & in_2 < 0 \end{cases}$$

$$\frac{\partial L}{\partial b_3} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial out_5} \frac{\partial out_3}{\partial in_3} \frac{\partial in_3}{\partial b_3} = \begin{cases} \left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) f(z)(1-f(z)), & in_3 > 0 \\ 0, & in_3 < 0 \end{cases}$$

$$\frac{\partial L}{\partial c} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial c} = \frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}} \cdot f(z)(1-f(z))$$