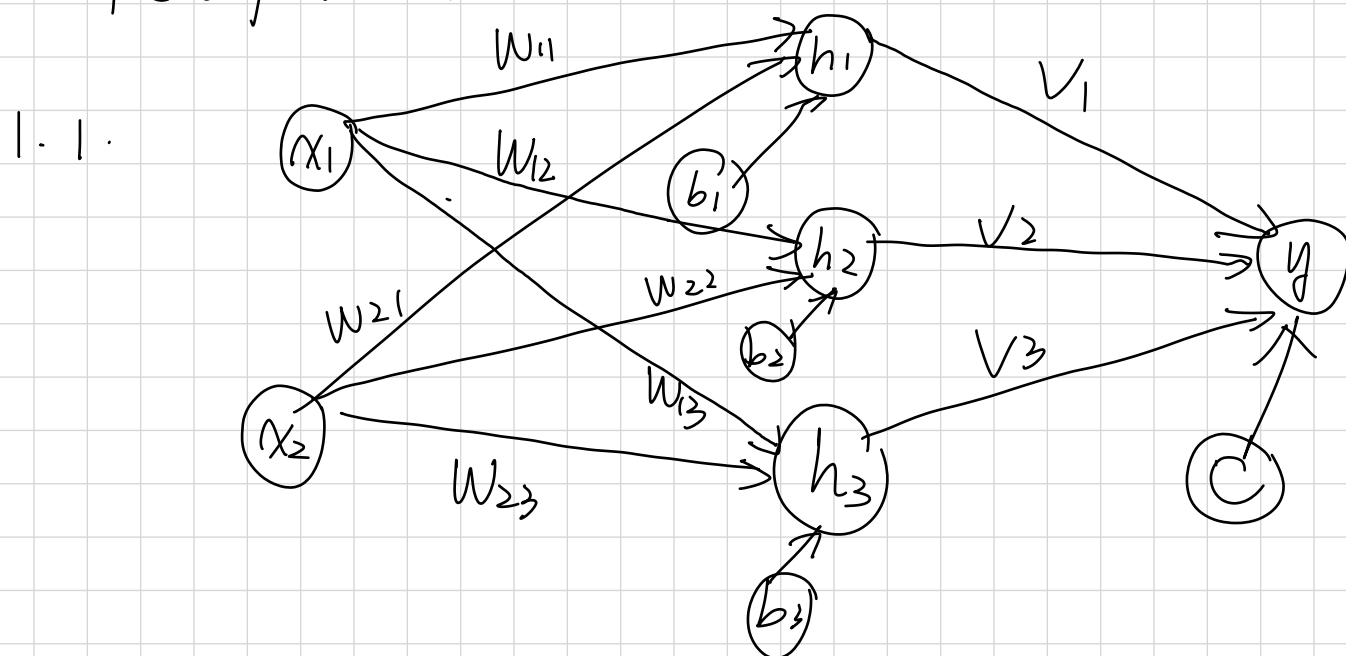# ANLY-590   Homework 2   Written Questions

Q1: Feedforward:

1.1.

1.2. This network is used for binary classification.

⟹ We use sigmoid function as the output function.

$$\Rightarrow \hat{y} = \sigma(V_1 h_1 + V_2 h_2 + V_3 h_3 + c)$$

$$h_1 = \max(W_{11} x_1 + W_{21} x_2 + b_1, 0)$$
$$h_2 = \max(W_{12} x_1 + W_{22} x_2 + b_2, 0)$$
$$h_3 = \max(W_{13} x_1 + W_{23} x_2 + b_3, 0)$$

Q2: Gradient Descent

2.1:
$$\frac{\partial f}{\partial x} = -3x^2 + 200x - 200y^2$$

$$\frac{\partial f}{\partial y} = 400y^3 - 400xy.$$

Q₃: Backprop

3.1. For sigmoid output function, we use binary crossentropy as loss function.

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

$$\Rightarrow \frac{dL}{dy} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right)$$

$$\frac{\partial L}{\partial V} = -\frac{dL}{dy} \cdot \frac{\partial y}{\partial V} = -\left(\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}\right) \cdot h \cdot \sigma(V^T h + c)(1 - \sigma(V^T h + c))$$

$$\frac{\partial L}{\partial W} = \frac{dL}{dy} \cdot \boxed{\frac{\partial y}{\partial h}} \cdot \boxed{\frac{\partial h}{\partial W}}$$

$$\rightarrow \frac{\partial y}{\partial h} = V^T \cdot \sigma(V^T h + c)(1 - \sigma(V^T h + c))$$

$$\rightarrow \frac{\partial h}{\partial W} = x^T \cdot \text{Relu}'(W^T x + b).$$

$$\frac{\partial L}{\partial b} = \frac{dL}{dy} \cdot \frac{\partial y}{\partial h} \cdot \boxed{\frac{\partial h}{\partial b}} \qquad \frac{\partial L}{\partial c} = \frac{dL}{dy} \cdot \boxed{\frac{\partial y}{\partial c}}$$

$$\frac{\partial h}{\partial b} = \text{Relu}'(W^T x + b) \qquad \frac{\partial y}{\partial c} = \sigma(V^T h + c) \cdot (1 - \sigma(V^T h + c))$$