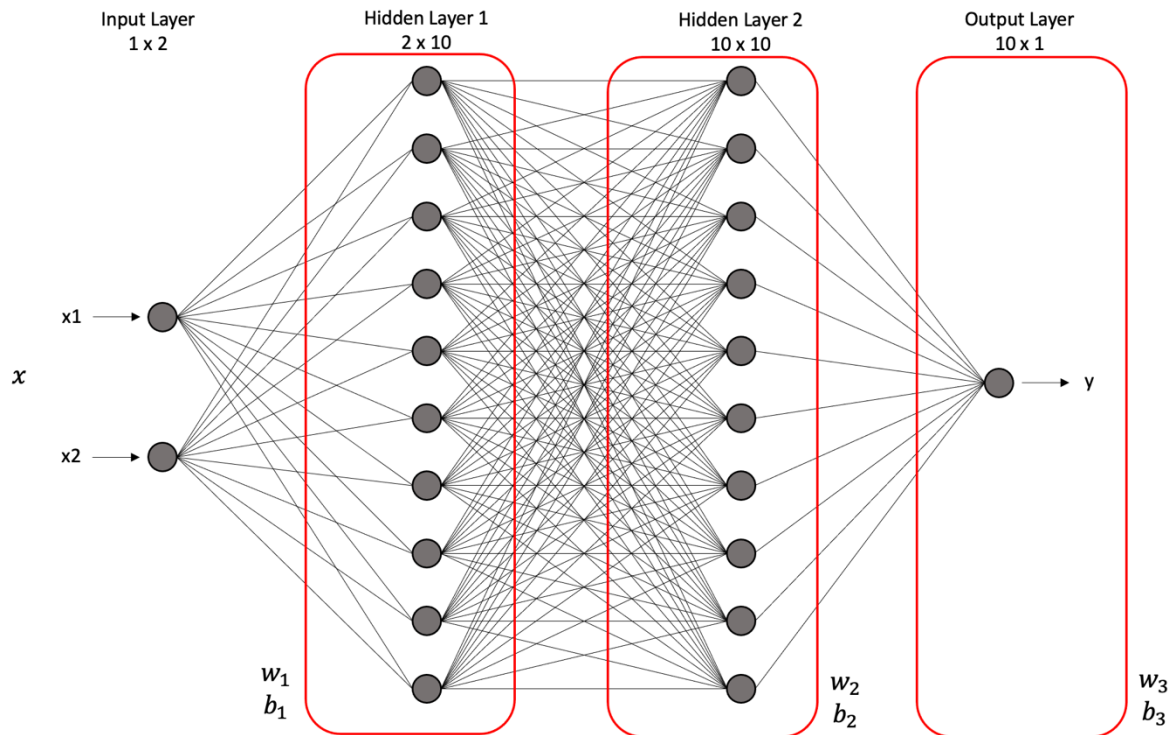


## BS6207 Homework 1

Isaac Lin

The figure below shows the neural network with 2 inputs, 10 x 10 hidden layers and 1 output.



The loss function is defined as follows:

$$\text{Loss} = \frac{1}{n} \sum_i (y_{\text{predict}} - y)^2 \quad \text{General equation}$$

$$= (y_{\text{predict}} - y)^2 \quad \text{Batch} = 1$$

I generated the labels using `torch.manual_seed(100)`.

$$x_1 = 0.1117, x_2 = 0.8158$$

$$y = \frac{x_1^2 + x_2^2}{2}$$

$$= \frac{0.1117^2 + 0.8158^2}{2}$$

$$= 0.3390$$

## Forward Propagation

$$h_1 = x \cdot w_1 + b_1$$

$$h_{1\_sigmoid} = \sigma(h_1)$$

$$h_2 = h_{1\_sigmoid} \cdot w_2 + b_2$$

$$h_{2\_sigmoid} = \sigma(h_2)$$

$$y = h_{2\_sigmoid} \cdot w_3 + b_3$$

## Backward Propagation

$$\frac{\delta L}{\delta w_3} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial w_3}$$

$$\frac{\delta L}{\delta b_3} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial b_3}$$

$$\frac{\delta L}{\delta w_2} = \frac{\partial L}{\partial h_2} \cdot \frac{\partial h_2}{\partial w_2} = \frac{\partial L}{\partial h_{2\_sigmoid}} \cdot \frac{\partial h_{2\_sigmoid}}{\partial h_2} \cdot \frac{\partial h_2}{\partial w_2}$$

$$\frac{\delta L}{\delta b_2} = \frac{\partial L}{\partial h_2} \cdot \frac{\partial h_2}{\partial b_2}$$

$$\frac{\delta L}{\delta w_1} = \frac{\partial L}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_1} = \frac{\partial L}{\partial h_{1\_sigmoid}} \cdot \frac{\partial h_{1\_sigmoid}}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_1}$$

$$\frac{\delta L}{\delta b_1} = \frac{\partial L}{\partial h_1} \cdot \frac{\partial h_1}{\partial b_1}$$

*torch\_autograd.dat* and *my\_autograd.dat* gives the same values of precision of up to 4 decimal places.