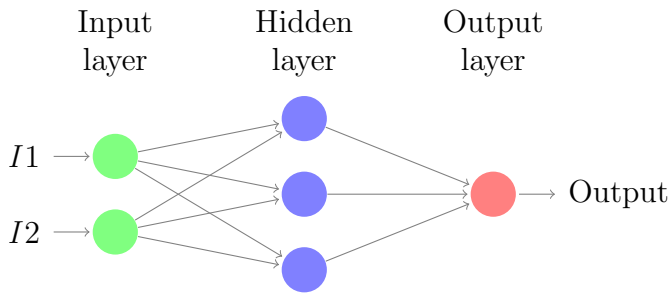
ANLY-590 Assignment 2
Yilun Zhu

1 Feedforward: Building a ReLU neural network

1.1 Draw a netwrok



1.2 Write out the mathematical equation for the output of this network

$$\hat{y} = \sigma(h_1 v_1 + h_2 v_2 + h_3 v_3 + c_1)$$
$$h_1 = ReLU(x_1 w_{11} + x_2 w_{21} + b_1)$$
$$h_2 = ReLU(x_1 w_{12} + x_2 w_{22} + b_2)$$
$$h_3 = ReLU(x_1 w_{13} + x_2 w_{23} + b_3)$$

2 Gradient Descent

2.1 What are the partial derivatives of f with respect to x and to y?

$$\frac{\partial f(x,y)}{\partial x} = -3x^2 - 200(y^2 - x)$$
$$\frac{\partial f(x,y)}{\partial y} = 400 * y * (y^2 - x)$$

3 Backprop

3.1 For the same network in Question 1, derive expressions of the gradient of the Loss function with respect to each of the model parameters

Suppose the Loss function is Binary Cross Entropy since it's a binary classification problem.

$$\frac{\partial L}{\partial V} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial out} \cdot \frac{\partial out}{\partial V} = \sum_{i=1}^{n} (\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}) \cdot \sigma'(out) \cdot h_i$$

$$\frac{\partial L}{\partial c} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial out} \cdot \frac{\partial out}{\partial c} = (\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}) \cdot \sigma'(out)$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial out} \cdot \frac{\partial out}{\partial H} \cdot \frac{\partial h_1}{\partial out'} \cdot \frac{\partial out'}{\partial W} = \sum_{i=1}^{n} \sum_{j=1}^{m} (\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}) \cdot \sigma'(out) \cdot v_i \cdot ReLU'(out'_i) \cdot x_j$$

$$\frac{\partial L}{\partial B} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial out} \cdot \frac{\partial out}{\partial H} \cdot \frac{\partial h_1}{\partial out'} \cdot \frac{\partial out'}{\partial B} = \sum_{i=1}^{n} (\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}}) \cdot \sigma'(out) \cdot v_i \cdot ReLU'(out'_i)$$