

# MACHINE LEARNING – WRITTEN ASSIGNMENT 3

Philip Hartout

November 7, 2016

1.

2.

3.

4. 1. With inputs  $x_1 = 0.5$  and  $x_2 = 0.9$ , we have the following values for the first and only hidden layer nodes:

$$a_1^{(2)} = \frac{1}{1 + e^{-(0.2+0.5 \cdot 0.5+0.5 \cdot 0.9)}} = 0.710949502625$$
$$a_2^{(2)} = \frac{1}{1 + e^{-(0.2+0.5 \cdot 0.1+0.9 \cdot 0.7)}} = 0.706822221094$$

Finally, we have the value for the output layer:

$$a_{\text{output}} = \frac{1}{1 + e^{0.2+0.710949502625+0.706822221094 \cdot 2}} = 0.910893519678$$

2.

$$\delta_1^{(3)} = a^{(3)} - y = 0.910893519678 - 1 = -0.089106480322$$

$$\delta_1^{(2)} = -0.089106480322 \cdot 0.710949502625 \cdot (1 - 0.710949502625) = -0.0183114090924$$

$$\delta_2^{(2)} = -0.089106480322 \cdot 2 \cdot 0.706822221094 \cdot (1 - 0.706822221094) = -0.036930103935$$

Now, since we have one output node, we update the weights without the regularisation term.

$$\Delta_{11}^{(2)} = 0.710949502625 \cdot (-0.0183114090924) = -0.0130184871866$$

$$\Delta_{21}^{(2)} = 0.706822221094 \cdot (-0.036930103935) = -0.0261030180886$$

Since  $\frac{\partial}{\partial \Theta_{ij}^{(l)}} J(\Theta) = D_{ij}^{(l)} = \Delta_{ij}^{(l)}$  in this case, we can update the weights as follows:

$$\Theta_{11}^{(2)} = 1 - (-0.0130184871866) = 1.01301848719$$

$$\Theta_{12}^{(2)} = 2 - (-0.0261030180886) = 2.02610301809$$

5. 1. The values of  $w_0 = -1$ ,  $w_1 = 1$  and  $w_2 = 1$  refer to the weights of the connections.  
2. a. We don't require a bias neuron.  $w_1 = 1$  and  $w_2 = -1$ .  
b. A XOR B.

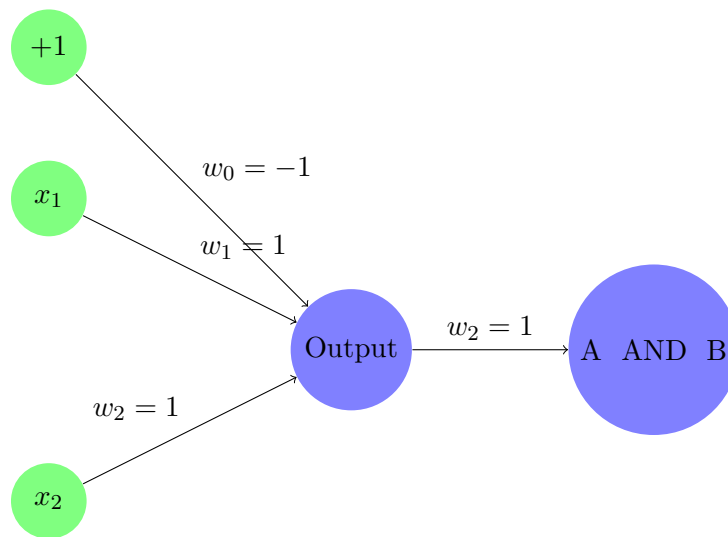


Figure 1:  $A \text{ AND } B$

