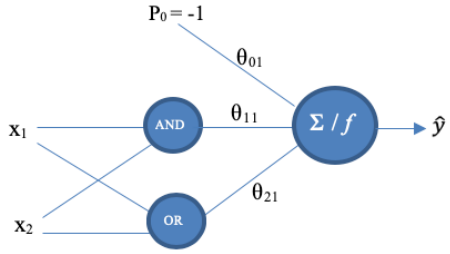
School of Computing and Information Systems
The University of Melbourne
COMP90049 Introduction to Machine Learning (Semester 1, 2023)
Workshop: week 10

1. Consider the two levels deep network illustrated below. It is composed of three perceptron. The two perceptron of the first level implement the AND and OR function, respectively.



Determine the weights θ_{11}, θ_{21} and bias θ_{01} such that the network implements the XOR function. The initial weights are set to zero, i.e., $\theta_{01} = \theta_{11} = \theta_{21} = 0$, and the learning rate η (eta) is set to 0.1.

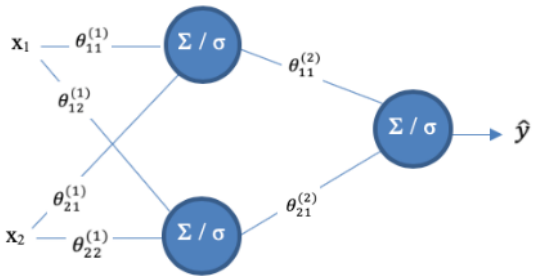
Notes:

- The input function for the perceptron on level 2 is the weighted sum (Σ) of its input.
- The activation function *f* for the perceptron on level 2 is a *step function*:

$$f = \begin{cases} 1 & if \sum > 0 \\ 0 & otherwise \end{cases}$$

- Assume that the weights for the perceptron of the first level are given.

2. Consider the following multilayer perceptron.



The network should implement the XOR function. Perform one epoch of *backpropagation* as introduced in the lecture on multilayer perceptrons.

Notes:

- The activation function *f* for a perceptron is the *sigmoid function*:

$$f(x) = \frac{1}{1 + e^{-x}}$$

- The thresholds are not shown in the network. The threshold nodes are set to -1.
- Use the following initial parameter values:

$$\theta_{01}^{(1)} = 2 \qquad \theta_{02}^{(1)} = -1 \qquad \theta_{01}^{(2)} = -2$$
$$\theta_{11}^{(1)} = 6 \qquad \theta_{12}^{(1)} = 8 \qquad \theta_{11}^{(2)} = 6$$
$$\theta_{21}^{(1)} = -6 \qquad \theta_{22}^{(1)} = -8 \qquad \theta_{21}^{(2)} = -6$$

- The learning rate is set to $\eta = 0.7$

i. Compute the activations of the hidden and output neurons.

ii. Compute the error of the network.

iii. Backpropagate the error to determine $\Delta\theta_{ij}$ for all weights θ_{ij} and updates the weight θ_{ij}.