

School of Computing and Information Systems  
The University of Melbourne  
COMP90049 Introduction to Machine Learning (Semester 1, 2021)  
Workshop: Week 7

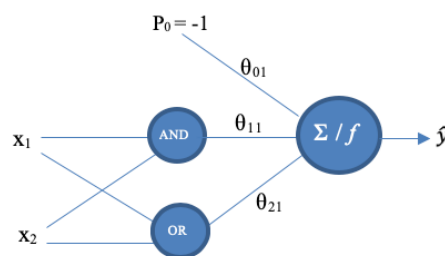
1. Why is a perceptron (which uses a **sigmoid** activation function) equivalent to *logistic regression*?
2. Consider the following training set:

$(x_1, x_2)$	$y$
(0,0)	0
(0,1)	1
(1,1)	1

With the bias value of 1, the initial weight function of  $\theta = \{\theta_0, \theta_1, \theta_2\} = \{0.2, -0.4, 0.1\}$  and learning rate of  $\eta = 0.2$ .

Consider the activation function of the perceptron as the step function  $f = \begin{cases} 1 & \text{if } \Sigma > 0 \\ 0 & \text{otherwise} \end{cases}$

- a) Can the perceptron learn a perfect solution for this data set?
  - b) Draw the perceptron graph and calculate the accuracy of the perceptron on the training data before training?
  - c) Using the perceptron *learning rule* and the learning rate of  $\eta = 0.2$ . Train the perceptron **for one epoch**. What are the weights after the training?
  - d) What is the accuracy of the perceptron on the training data after training for one epoch? Did the accuracy improve?
3. 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  $\theta_{11}$ ,  $\theta_{21}$  and bias  $\theta_{01}$  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  $\eta$  (eta) is set to 0.1.

Notes:

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

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

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