National University of Singapore
School of Computing
CS3244 Machine Learning

Tutorial 4: Neural Networks

Issue: February 17, 2022 Due: Feb 28, 2022

Important Instructions:

- *Your solutions for this tutorial must be TYPE-WRITTEN.*
- *Make TWO copies of your solutions: one for you and one to be SUBMITTED TO THE TUTOR IN CLASS. Your submission in your respective tutorial class will be used to indicate your CLASS ATTENDANCE. Late submission will NOT be entertained.*
- *Indicate your NAME, STUDENT NUMBER, and TUTORIAL GROUP in your submitted solution.*
- *YOUR SOLUTION TO QUESTION* **BL 7** *will be GRADED for this tutorial.*
- *You may discuss the content of the questions with your classmates. But everyone should work out and write up ALL the solutions by yourself.*

BL 6 Supposing the weights w_1 and w_2 of a perceptron (see page 6 of "Neural Networks" lecture slides) are both set to the value of **0.5**, derive the **largest** possible range of the values of w_0 that can be set for the perceptron to represent the AND gate (i.e., AND(x_1, x_2)). Assume that the inputs x_1 and x_2 and output $o(x_1, x_2)$ of the perceptron are Boolean with the values of **1 or –1**.
Show the steps of your derivation. **No marks will be awarded for not doing so.**

Solution. $-1 < w_0 \leq 0$.

BL 7 (Final Exam AY2017/18) Supposing the weights w_1 and w_2 of a perceptron (see page 6 of "Neural Networks" lecture slides) are both set to the value of **–1**, derive the **largest** possible range of the values of w_0 that can be set for the perceptron to represent the NAND gate (i.e., NAND(x_1, x_2)). Assume that the inputs x_1 and x_2 and output $o(x_1, x_2)$ of the perceptron are Boolean with the values of **1 or –1**.
Show the steps of your derivation. **No marks will be awarded for not doing so.**

Solution. $0 < w_0 \leq 2$.

BL 8 In every multi-layer network of perceptron units described below,

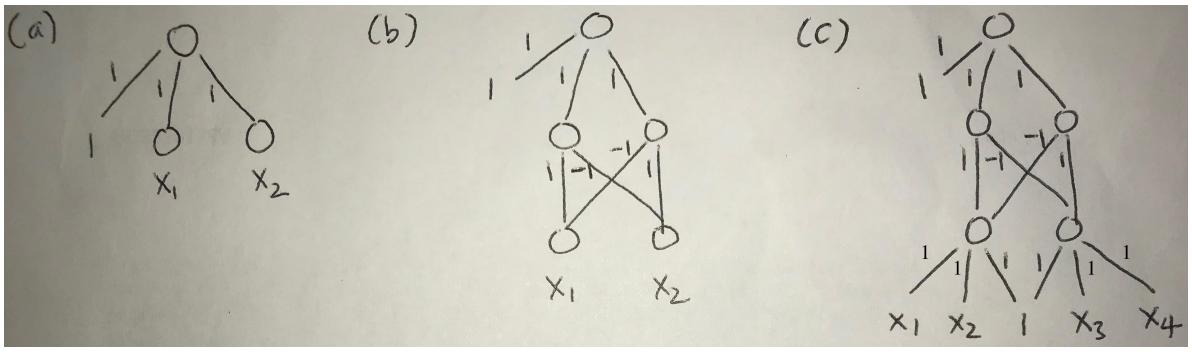
- There should be only one (Boolean) output node and an input node for every (Boolean) input.
- A Boolean is -1 if false, and 1 if true.
- The activation function of every (non-input) node is a -1 to 1 step function, including that of the output node (see page 6 of the "Neural Networks" lecture slides).
- Your weights must be **integers** and kept **small**, but possibly negative.
- Keep your networks as symmetric as possible – doing this in questions a and b may help you in question c.
- You don't have to draw edges with weight 0.

(a) Construct and draw a perceptron network with no hidden layers that implements (x_1 OR x_2).

(b) Construct and draw a perceptron network with one hidden layer (with two nodes) that implements (x_1 XOR x_2).

(c) Combining your constructions in questions a and b, construct and draw a perceptron network with two hidden layers (two nodes each) that implements (x_1 OR x_2) XOR (x_3 OR x_4).

Solution.



TM 4.1 What are the values of weights w_0, w_1, and w_2 for the perceptron whose decision surface is illustrated in figure a on page 7 of the "Neural Networks" lecture slides? Assume the surface crosses the x_1 axis at -1, and the x_2 axis at 2.

Solution. $w_0 = w_1 = -2w_2$, $w_0, w_1 < 0$ and $w_2 > 0$. For example, $w_0 = -2$, $w_1 = -2$, and $w_2 = 1$.

TM 4.3 Consider two perceptrons defined by the threshold expression $w_0 + w_1x_1 + w_2x_2 > 0$. Perceptron A has weight values $w_0 = 1$, $w_1 = 2$, $w_2 = 1$ and perceptron B has the weight values $w_0 = 0$, $w_1 = 2$, $w_2 = 1$. True or false? Perceptron A is *more general than* perceptron B; the notion of *more general than* is previously defined on page 12 of the "Concept Learning" lecture slides.

Solution. True.