The University of Adelaide
School of Computer Science

Artificial Intelligence, 2019
Worksheet 3

Question 1 (Question 20.21 in AIMA 2ed)

20.21 The main purpose of this exercise is to make concrete the notion of the *capacity* of a function class (in this case, linear halfspaces). It can be hard to internalize this concept, but the examples really help.

a. Three points in general position on a plane form a triangle. Any subset of the points can be separated from the rest by a line, as can be seen from the two examples in Figure S20.5(a).

b. Figure S20.5(b) shows two cases where the positive and negative examples cannot be separated by a line.

c. Four points in general position on a plane form a tetrahedron. Any subset of the points can be separated from the rest by a plane, as can be seen from the two examples in Figure S20.5(c).

d. Figure S20.5(d) shows a case where a negative point is inside the tetrahedron formed by four positive points; clearly no plane can separate the two sets.

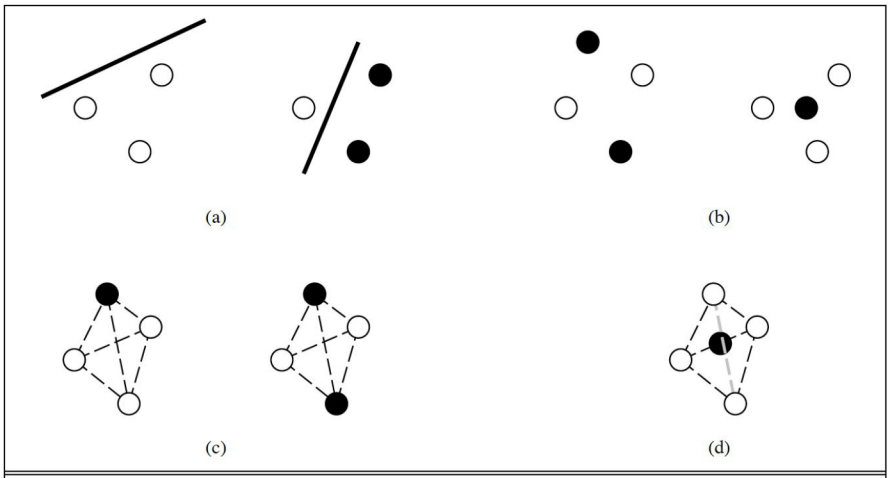

Figure S20.5 Illustrative examples for Ex. 20.21.

Question 2 (Question 20.11 in AIMA 2ed)

20.11 XOR (in fact any Boolean function) is easiest to construct using step-function units. Because XOR is not linearly separable, we will need a hidden layer. It turns out that just one hidden node suffices. To design the network, we can think of the XOR function as OR with the AND case (both inputs on) ruled out. Thus the hidden layer computes AND, while the output layer computes OR but weights the output of the hidden node negatively. The network shown in Figure S20.3 does the trick.

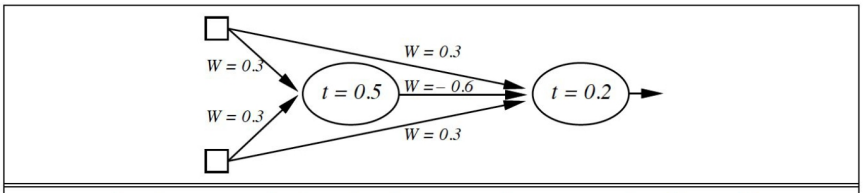

Figure S20.3 A network of step-function neurons that computes the XOR function.

Question 3 (Question 20.19 in AIMA 2ed)

The Error under $L2$ loss is given by:

$$E = \frac{1}{2}\sum_i (y_i - a_1)^2 = \frac{1}{2}[80(1-a_1)^2 + 20(0-a_1)^2] = 50a_1^2 - 80a_1 + 50$$

Derivative of error w.r.t. a_1 is then:

$$\frac{\partial E}{\partial a_1} = 100a_1 - 80$$

Setting derivative equal to zero (i.e. minimum value) yields $a_1 = 0.8$.

Question 4

Figure and calcs below use Heaviside step function $H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$ which is the derivative of Relu. Also note that derivative of $l_1(x)$ loss is $2H(x) - 1$