Machine Learning 2016/2017: Assignment 3

Elise Mol

3. We have a Neural Network with 1 hidden layer. Input and hidden layer both have 2 nodes. There is 1 output node. The values of theta for bias nodes are 0.2. The vector θ(1) for layer 1 is [0.5, 0.1, 0.5, 0.7] and θ(2) for layer 2 is [1, 2].
3.1 Calculate by hand the activations of all nodes for x1 = 0.5 and x2 = 0.9.

We take θ_{11}^{1} to be 0.5, θ_{21}^{1} to be 0.1, θ_{12}^{1} to be 0.5 and θ_{22}^{1} to be 0.7 for layer one. For layer two we take θ_{11}^{2} to be 1 and θ_{12}^{2} to be 2.

$a_l^j = g(\Sigma_{i=1}^{i} \Theta^{j-1} * x_l^{j-1})$
$g = 1 / 1 + e^{-z}$

$z_1^2 = \theta_{10}^{1} * 1 + \theta_{11}^{1} * x_1 + \theta_{12}^{1} * x_2 = 0.2 * 1 + 0.5 * 0.5 + 0.5 * 0.9 = 0.9$
$z_2^2 = \theta_{20}^{1} * 1 + \theta_{21}^{1} * x_1 + \theta_{22}^{1} * x_2 = 0.2 * 1 + 0.1 * 0.5 + 0.7 * 0.9 = 0.88$
$z_1^3 = \theta_{10}^{2} * 1 + \theta_{11}^{2} * a_1^2 + \theta_{12}^{2} * a_2^2$

$a_1^2 = 1 / 1 + e^{-z12} = 1 / 1 + e^{-0.9} = 0.7109495$
$a_2^2 = 1 / 1 + e^{-z22} = 1 / 1 + e^{-0.88} = 0.7068222$

$z_1^3 = 0.2 * 1 + 1 * 0.7109495 + 2 * 0.7068222 = 2.3245939$

$a_1^3 = 1 / 1 + e^{-2.3245939} = 0.910893516$

The activations of all nodes are
$a_1^1 = x1 = 0.5$
$a_2^1 = x2 = 0.9$
$a_1^2 = 0.7109495$
$a_2^2 = 0.7068222$
$a_1^3 = 0.910893516$

3.2 Suppose the correct output is 1. Calculate the errors for all nodes and the updates of the weights (for 1 iteration).

$\delta_1^3 = a_1^3 - y_1 = 1 - 0.910893516 - 1 = 0.089106484$

$\delta_2^2 = (\theta_{12}^2 * \delta_1^3) *^0 g'(z_2^2) = 2 * 0.089106484 *^0 g'(z_2^2)$
$g'(z_2^2) = a_2^2 *^0 (1 - a_2^2) = 0.7068222 *^0 (1 - 0.7068222)$
$\delta_2^2 = 2 * 0.089106484 *^0 (0.7068222 *^0 (1 - 0.7068222))$

$\delta_1^2 = (\theta_{11}^2 * \delta_1^3) *^0 g'(z_1^2) = (1 * 0.089106484) *^0 g'(z_1^2)$
$g'(z_1^2) = a_1^2 *^0 (1 - a_1^2) = 0.7109495 *^0 (1 - 0.7109495)$
$\delta_1^2 = (1 * 0.089106484) *^0 (0.7109495 *^0 (1 - 0.7109495))$

$\delta_2^1 = (\theta_{12}^1 * \delta_1^2 + \theta_{22}^1 * \delta_2^2) *^0 g'(z_2^1) = (0.5 * \delta_1^2 + 0.7 * \delta_2^2) * g'(z_2^1)$
$g'(z_2^1) = a_2^1 *^0 (1 - a_2^1) = 0.7109495 *^0 (1 - 0.7109495)$

$\delta_1^1 = (\theta_{11}^1 * \delta_1^2 + \theta_{21}^1 * \delta_2^2) *^0 g'(z_1^1) = (0.5 * \delta_1^2 + 0.5 * \delta_2^2) * g'(z_2^1)$
$g'(z_2^1) = a_2^1 *^0 (1 - a_2^1) = 0.6224593 *^0 (1 - 0.6224593)$

The errors for all nodes
$\delta_1^3 = 0.089106484$
$\delta_2^2 = 2 * 0.089106484 *^0 (0.7068222 *^0 (1 - 0.7068222)) = 0.036930107$
$\delta_1^2 = (1 * 0.089106484) *^0 (0.7109495 *^0 (1 - 0.7109495)) = 0.0183114099$
$\delta_2^1 = (0.5 * \delta_1^2 + 0.7 * \delta_2^2) *^0 (0.7109495 *^0 (1 - 0.7109495)) = 0.007193904$
$\delta_1^1 = (0.5 * \delta_1^2 + 0.5 * \delta_2^2) *^0 (0.6224593 *^0 (1 - 0.6224593)) = 0.00649098$

4. In this exercise you will use a perceptron (instead of sigmoid function), that is explained in the slides of Mitchell. You can find those slides here.
Please refer to pages 78 and 79.
4.1) What are the values of weights w0, w1 and w2 for the perceptron whose decision surface is illustrated in figure b on page 79? Assume the surface crosses the x_1 axis at -1 and the x_2 axis at 2.

Because a decision surface in 2D is always a line we simply need to represent this line with the formula: $w_2 x_2 = w_0 x_0 + w_1 x_1$
With the two points given to us, (0,2) and (-1,0), and the knowlegde that x0 is a bias term which usually is 1, we can determine that $w_1 = 2$, $w_2 = -1$ *and* $w_0 = -2$.

(a) Design a two input perceptron that implements the Boolean function A AND (NOT B).

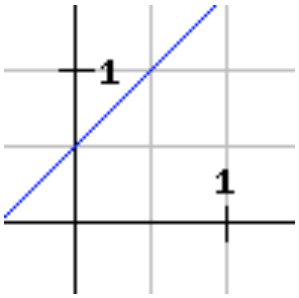
To design the input perceptron that used and and not we need to define those first.

NOT	
in	*out*
1	*0*
0	*1*

AND		
in1	*in2*	*out*
0	*0*	*0*
0	*1*	*0*
1	*0*	*0*
1	*1*	*1*

So our the inplementation of the Boolean function A AND (NOT B) of our two input perceptron is:

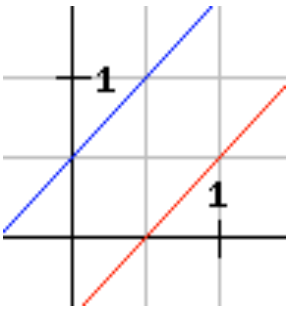
A	*AND*	*(NOT*	*B)*
1	*0*	*0*	*1*
1	*1*	*1*	*0*
0	*0*	*0*	*1*
0	*0*	*1*	*0*



The two input perceptron could for example be the graph on the left, where we have the points (0, 0) (0, 1) and (1, 1) on the right of the line, and point (1, 0) on the right. The x-axis is A and the y-axis is B.

(b) Design a two layer network of perceptrons that that implements A XOR B

A	*XOR*	*B*
0	*0*	*0*
0	*1*	*1*
1	*1*	*0*
1	*0*	*1*



Where the points (0, 0) and (1, 1) are inbetween the two lines, and (0, 1) and (1, 0) are outside the lines. The x-axis is A and the y-axis is B.