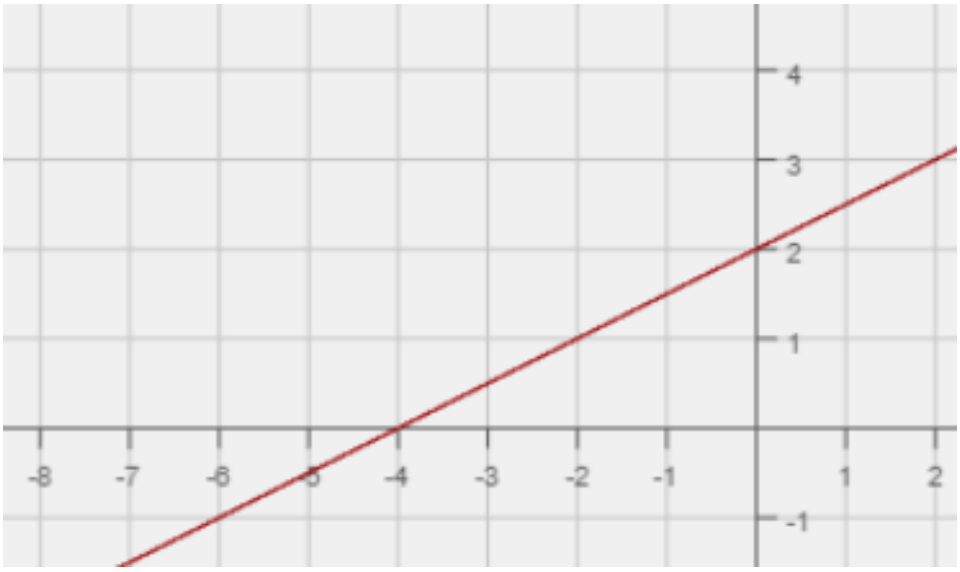
Isabel Walter
Written Assignment 3
Machine Learning
Worked together with Carine Candel

1.
0 = 4 * x0 + x1 -2 * x2
0 = 4 + x1 -2 * x2
2 * x2 = 4 + x1
x2 = 2 + 0.5 * x1



4 * x0 + x1 -2 * x2 gives -1 above the line and 1 under the line.

2.
If you have no parameters then your hypothesis will be a horizontal line → underfitting.
If you have as many parameters as you have examples your hypothesis will fit your training set perfectly but it won't do so well on the test set → overfitting.

3.
1
$\theta_{10}^{(1)} = 0.2$ $\theta_{10}^{(2)} = 0.2$
$\theta_{20}^{(1)} = 0.2$ $\theta_{11}^{(2)} = 1$
$\theta_{11}^{(1)} = 0.5$ $\theta_{12}^{(2)} = 2$
$\theta_{12}^{(1)} = 0.1$
$\theta_{21}^{(1)} = 0.5$
$\theta_{22}^{(1)} = 0.7$

$a_1^{(1)} = x = 0.5$
$a_2^{(1)} = x = 0.9$
$z_1^{(2)} = \theta_{10}^{(1)} * 1 + \theta_{11}^{(1)} * a_1^{(1)} + \theta_{12}^{(1)} * a_2^{(1)} = 0.2 + 0.25 + 0.09 = 0.54$
$z_2^{(2)} = \theta_{20}^{(1)} * 1 + \theta_{21}^{(1)} * a_1^{(1)} + \theta_{22}^{(1)} * a_2^{(1)} = 1.08$

$a_1^{(2)} = \frac{1}{1+e^{-0.54}}$ = 0.6318124177 => 0.63
$a_2^{(2)} = \frac{1}{1+e^{-1.08}}$ = 0.7464939833 => 0.75
$z_1^{(3)} = \theta_{10}^{(2)} * 1 + \theta_{11}^{(2)} * a_1^{(2)} + \theta_{12}^{(2)} * a_2^{(2)}$ = 2.3248
$a_1^{(3)} = \frac{1}{1+e^{-2.3248}}$ = 0.9109102742 => 0.91

2
$\delta_1^{(3)} = a_1^{(3)} - y$ = 0.9109102742 – 1 = - 0.0890897258 => -0.09
$\delta_1^{(2)} = \theta_{11}^{(2)} * \delta_1^{(3)} * a_1^{(2)} * (1 - a_1^{(2)})$ = - 0.0207245408 => -0.02
$\delta_2^{(2)} = \theta_{12}^{(2)} * \delta_1^{(3)} * a_2^{(2)} * (1 - a_2^{(2)})$ = - 0.033718807 => -0.03

1 iteration:
α = 0.1
$\theta_{11}^{(2)} = \theta_{11}^{(2)} - \alpha * a_1^{(2)} * \delta_1^{(3)}$ = 1.0056288
$\theta_{12}^{(2)} = \theta_{12}^{(2)} - \alpha * a_2^{(2)} * \delta_1^{(3)}$ = 2.006650494
$\theta_{11}^{(1)} = \theta_{11}^{(1)} - \alpha * a_1^{(1)} * \delta_1^{(2)} =$ 0.501036227
$\theta_{12}^{(1)} = \theta_{12}^{(1)} - \alpha * a_2^{(1)} * \delta_1^{(2)}$ = 0.1018652087
$\theta_{21}^{(1)} = \theta_{21}^{(1)} - \alpha * a_1^{(1)} * \delta_2^{(2)}$ = 0.5016859404
$\theta_{22}^{(1)} = \theta_{22}^{(1)} - \alpha * a_2^{(1)} * \delta_2^{(2)}$ = 0.9030346926

4.
1
Equation of the line:
w0 + w1x1 + w2x2 = 0
Given points:
P(-1,0) and Q(0,2)
w0 - w1 = 0
w0 + 2 * w2 = 0
w0 = w1 = -2 * w2
e.g. w0 and w1 are 2 and w2 is -1
If you fill this in for, for example, a point on the right of the line with x1 = 1 and x2 = 0 then
w0 + w1x1 + w2x2 = 2 + 2 * 1 + 0 = 4 is bigger than 0 → would give an outcome of 1 but is supposed to be -1.

So w0 and w1 are -2 and w2 is 1.

2
A

A	B	A ^ not B
-1	-1	-1
-1	1	-1
1	-1	1
1	1	-1

So line crosses the A axis at 1 and the B axis at -1.
w0 + w1x1 + w2x2 = 0
P(1, 0) and P(0,-1)

w0 + w1 = 0

w0 – w2 = 0
w0 = -w1 = w2

So e.g. w0 = 1 w1 = -1 and w2 = 1
If you fill this in for, for example, a point on the right of the line with x1 = 1 and x2 = -1 then
w0 + w1x1 + w2x2 = 1 – 1 – 1 = -1 which is smaller than 0 → outcome of -1 but is supposed to be 1 → negate weights:
Gives w0 = -1, w1 = 1, w2 = -1

B
A XOR B = (A∧¬B) ∨(¬A∧B)

A	B	P1 (A∧¬B)	P2 (¬A∧B)	P3 (A∧¬B)∨(¬A∧B)
-1	-1	-1	-1	-1
-1	1	-1	1	1
1	-1	1	-1	1
1	1	-1	-1	-1

For P1: w0 = -1, w1 = 1, w2 = -1
For P2:
Line crosses the A axis at -1 and the B axis at 1.
w0 + w1x1 + w2x2 = 0
P(-1, 0) and P(0,1)

w0 - w1 = 0
w0 + w2 = 0
w0 = w1 = -w2

So e.g. w0 = 1 w1 = 1 and w2 = -1
If you fill this in for, for example, a point on the right of the line with x1 = -1 and x2 = 1 then
w0 + w1x1 + w2x2 = 1 – 1 – 1 = -1 which is smaller than 0 → outcome of -1 but is supposed to be 1 → negate weights:
Gives w0 = -1, w1 = -1, w2 = 1

For P3:
w0 + w1x1 + w2x2 = 0
P(1, -1) or P(-1,1)
w0 + w1 – w2 = 0
w0 – w1 + w2 = 0
w0 = - w1 + w2 = w1 – w2
- w1 + w2 = w1 – w2
w2 = 2 * w1 – w2
2 * w2 = 2 * w1
w2 = w1 = w0
Take all the same weights, e.g. take for all weight 0.5.
If x1 = 1 and x2 = 1 then the perceptron should return 1.
w0 + w1x1 + w2x2 = 0.5 + 0.5 + 0.5 = 1.5, is bigger than 0 so perceptron would return 1.
Gives w0 = 0.5, w1 = 0.5 and w2 = 0.5.