Scalable Data Engineering, Exercise 8

Henry Haustein

Task 1

(a) False. Supervised Learning needs labelled data.

(b) False. The output is $f(Xw)$.

(c) True.

(d) False. For non-linearity you need at least one non-linear activation function.

(e) True.

(f) False. Three derivatives.

(g) False. Exploding gradients lead to instability because of huge changes in weights.

(h) False. Problem are dead neurons.

(i) True.

Task 2

Using an Excel Spreadsheet we get

x_1	x_2	t	w_{11}^{old}	w_{21}^{old}	z	a	w_{11}^{new}	w_{21}^{new}
1	0	1	-0,2	0,4	-0,2	0	0,1	0,4
0	1	1	0,1	0,4	0,4	0	0,1	0,7
0	0	0	0,1	0,7	0	0	0,1	0,7
1	1	1	0,1	0,7	0,8	0	0,4	1
1	0	1	0,4	1	0,4	0	0,7	1
0	1	1	0,7	1	1	1	0,7	1
0	0	0	0,7	1	0	0	0,7	1
1	1	1	0,7	1	1,7	1	0,7	1
1	0	1	0,7	1	0,7	0	1	1
0	1	1	1	1	1	1	1	1
0	0	0	1	1	0	0	1	1
1	1	1	1	1	2	1	1	1

1	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1
0	0	0	1	1	0	0	1	1
1	1	1	1	1	2	1	1	1

The weights converge to 1.

Task 3

Since we know the true values t_{a_1}, t_{a_2} and t_{a_3} we can train all three perceptrons individually. For the first perceptron we see that it should only fire if and only if x_1 fires. This leads to $w_{11} = 1$ and $w_{21} = -1$. For the second perceptron it is the opposite: It should only fire if and only if x_2 fires. Therefore $w_{21} = -1$ and $w_{22} = 1$.

The third perceptron models the XOR function which is not linear separable. Therefore we won't find weights w_{13} and w_{23} that will produce correct results.