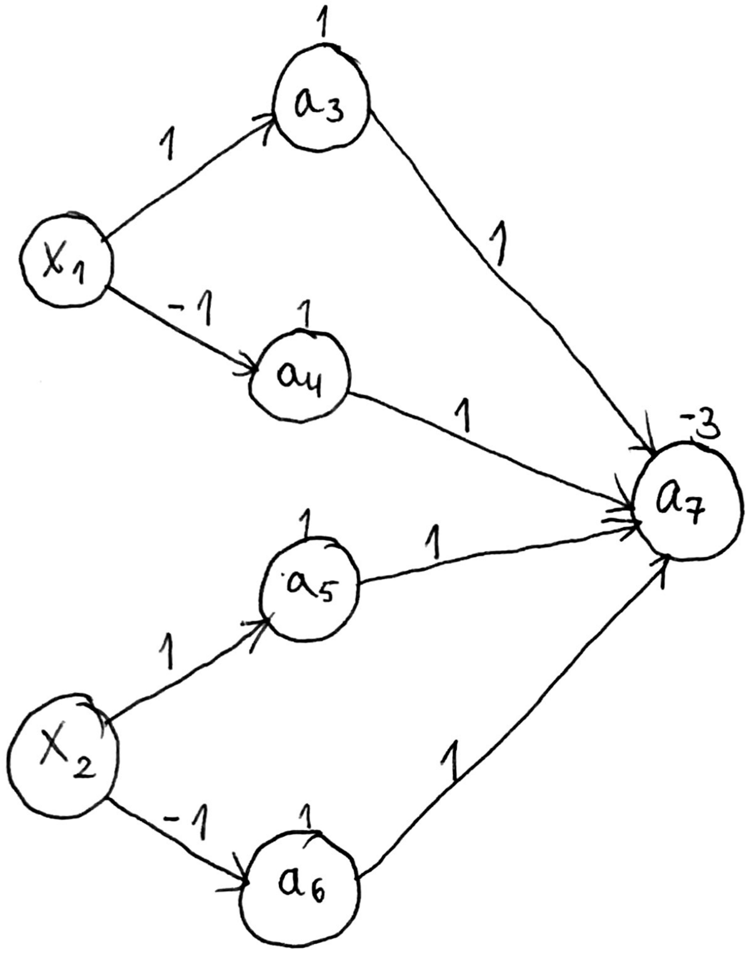
Machine Learning 1 WS18/19

Submission for Exercise Sheet 11

Michael Hoppe 362514
Wai Tang Victor Chan 406094
Jonas Piotrowski 399334
Aki Saksala 399293
Sourabh Raj 397371
Farnoush Rezaei Jafari 398708

Exercise Sheet 11

(1)



$w_{13} = 1$ $\quad b_3 = 1$

$w_{14} = -1$ $\quad b_4 = 1$

$w_{25} = 1$ $\quad b_5 = 1$

$w_{26} = -1$

$w_{37} = 1$

$w_{47} = 1$ $\quad b_6 = 1$

$w_{57} = 1$

$w_{67} = 1$ $\quad b_7 = -3$

assuming that the boundary is exclusive, i.e. $x_1 = 1$, $x_2 = 1$ belongs to the outer class B.

Exercise 2: Backward Propagation

$$E(w) = \| y(x, w) - t \|^2 = \left[\frac{1}{2}(a_8 - t_8)^2 + \frac{1}{2}(a_9 - t_9)^2 \right]$$

$$\frac{\partial E}{\partial w_{14}} = \frac{\partial E}{\partial Z_4} \cdot \frac{\partial Z_4}{\partial w_{14}} = \underbrace{\delta_4}_{\mathrm{I}} \cdot \underbrace{\frac{\partial Z_4}{\partial w_{14}}}_{\mathrm{II}}$$

$$\mathrm{I} \Rightarrow \frac{\partial Z_4}{\partial w_{14}} = \frac{\partial}{\partial w_{14}} \sum w_{i4} \cdot a_i = \frac{\partial w_{14} \cdot a_1}{\partial w_{14}} = a_1$$

$$\mathrm{II} \Rightarrow \delta_4 = \frac{\partial E}{\partial Z_4} = \frac{\partial E}{\partial Z_8} \cdot \frac{\partial Z_8}{\partial Z_4} + \frac{\partial E}{\partial Z_6} \cdot \frac{\partial Z_6}{\partial Z_4} = \underbrace{\delta_8}_{\mathrm{III}} \cdot \frac{\partial Z_8}{\partial a_4} \cdot \frac{\partial a_4}{\partial Z_4} + \underbrace{\delta_6}_{\mathrm{IV}} \cdot \frac{\partial Z_6}{\partial a_4} \cdot \frac{\partial a_4}{\partial Z_4}$$

$$\mathrm{III} \rightarrow \delta_8 = \frac{\partial E}{\partial Z_8} = \frac{\partial E}{\partial a_8} \cdot \frac{\partial a_8}{\partial Z_8} = (a_8 - t_8) \cdot (a_8 \cdot (1 - a_8))$$

like δ_4

$$\mathrm{IV} \quad \delta_6 = \delta_8 \cdot \frac{\partial Z_8}{\partial a_6} \cdot \frac{\partial a_6}{\partial Z_6} + \underbrace{\delta_9}_{\mathrm{V}} \cdot \frac{\partial Z_9}{\partial a_6} \cdot \frac{\partial a_6}{\partial Z_6}$$

$$\mathrm{V} \Rightarrow \delta_9 = (a_9 - t_9) \cdot (a_9 - (1 - a_9))$$

Then we conclude that:

$$\frac{\partial E}{\partial w_{14}} = \delta_4 \cdot \frac{\partial Z_4}{\partial w_{14}} = \left(\delta_8 \cdot \frac{\partial Z_8}{\partial a_4} \cdot \frac{\partial a_4}{\partial Z_4} + \delta_6 \cdot \frac{\partial Z_6}{\partial a_4} \cdot \frac{\partial a_4}{\partial Z_4} \right) a_1$$

$$\frac{\partial E}{\partial w_{16}} = \frac{\partial E}{\partial z_6} \cdot \frac{\partial z_6}{\partial w_{16}} = \underbrace{\delta_6} \cdot \underbrace{\frac{\partial z_6}{\partial w_{16}}}_{\mathrm{I}}$$

It is cakulated before

$$z_6 = x_1 \cdot w_{16} + a_4 \cdot w_{46}$$

$$\mathrm{I} \Rightarrow \frac{\partial z_6}{\partial w_{16}} = \frac{\partial (x_1 \cdot w_{16} + a_1 \cdot w_{46})}{\partial w_{16}} = x_1$$

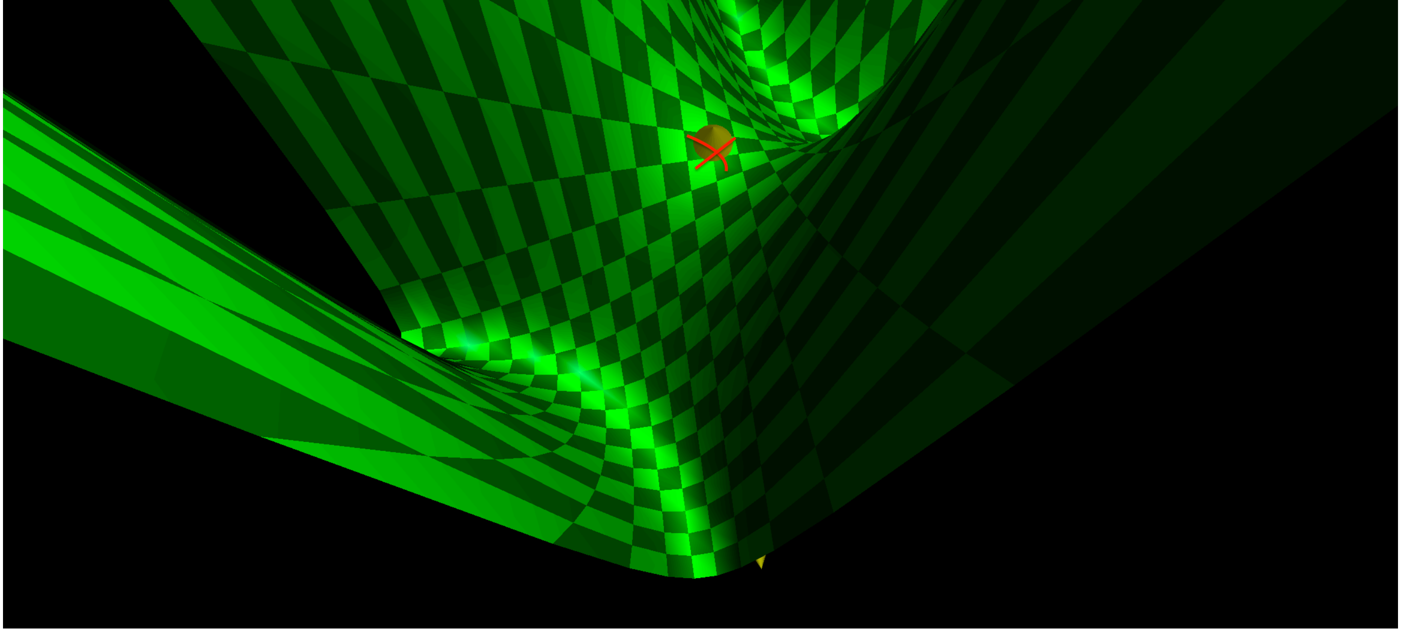
1. $(\omega_{12}, \omega_{23}) = (0, 0)$

$$E(\omega_{12}, \omega_{23}) = \sum_{n=1}^{3} \left(f(x^n) - t^n \right)^2$$

x	t	y	e
-1	-1	0	1
0	0	0	0
1	1	0	1
			2

for $n = 1 \to 3$

clearly E is not the minimumd.



Since the gradient is zero, the parameters would not move away from (0,0) in a simple gradient descent.

8.

(2)

$$\text{output} = x_1 w_{12} w_{23}$$

$$E(w_{12}, w_{23}) = 0$$

$$\Rightarrow \sum_{n=1}^{3} \left(t(x_i) - (t)^n \right)^2 = 0$$

$$\Rightarrow \left(-1 \cdot w_{12} w_{23} + 1 \right)^2 + \left(0 \cdot w_{12} w_{23} + 0 \right)^2 + \left(1 \cdot w_{12} w_{23} - 1 \right)^2 = 0$$

$$\Rightarrow 2 \left(w_{12} w_{23} - 1 \right)^2 = 0$$

$$\Rightarrow \boxed{w_{12} = \frac{1}{w_{23}}}$$

so, this implies.

$$\boxed{\text{output} = x_1}$$

output directly and only depends on input and so, can have infinite number of solutions.