Exercise 1:

$$E(W) = \frac{1}{N} \sum_n \| \vec{g}(\vec{s}^{(1)}_{(n)}, W) - \vec{l}^{(n)} \|^2$$

$$= \frac{1}{N} \sum_n \sum_h (S(\sum_i W_{hi} s_i^{(1)}) - l_h^{(n)})^2$$

For the derivative of $E(W)$ holds: (lecture video Lec11-Vid04)

$$\frac{d}{dW_{hi}} E(W) = \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \, S'(\sum_i W_{hi} s_i^{(1)}) \, s_i^{(1)}$$

1) $g_h(\vec{s}^{(1)}, W) = \sum_i W_{hi} s_i^{(1)} \Rightarrow S(x) = x$ with $x = \sum_i W_{hi} s_i^{(1)}$ and S is the activation function

$$\Rightarrow \frac{d}{dW_{hi}} E(W) = \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \underbrace{\frac{\partial}{\partial x} S(x)}_{=1} s_i^{(1)}$$

$$= \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \cdot s_i^{(1)}$$

$$\Rightarrow \Delta W_{hi} = -\varepsilon \cdot \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \cdot s_i^{(1)}$$

2) $g_h(\vec{s}^{(1)}, W) = \dfrac{\exp(2 \sum_i W_{hi} s_i^{(1)}) - 1}{\exp(2 \sum_i W_{hi} s_i^{(1)}) + 1} \Rightarrow S(x) = \dfrac{\exp(2x) - 1}{\exp(2x) + 1}$

$$\Rightarrow \underset{\substack{\uparrow \\ \text{Quotient rule}}}{S'(x)} = \frac{2e^{2x}(e^{2x}+1) - (e^{2x}-1)2e^{2x}}{(e^{2x}+1)^2} = \frac{4e^{2x}}{(e^{2x}+1)^2} = \frac{4}{(e^x + e^{-x})^2} = \frac{1}{\cosh^2(x)} = \operatorname{sech}^2(x)$$

$$\Rightarrow \frac{dE(W)}{dW_{hi}} = \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \cdot \operatorname{sech}^2 \left(\sum_i W_{hi} s_i^{(1)} \right) \cdot s_i^{(1)}$$

$$\Rightarrow \Delta W_{hi} = -\varepsilon \frac{1}{N} \sum_n 2 (s_h^{(0)} - l_h^{(n)}) \cdot \operatorname{sech}^2 \left(\sum_i W_{hi} s_i^{(1)} \right) \cdot s_i^{(1)}$$