

5) evaluate the gradients

$$\frac{\partial E}{\partial b_i^{(e)}} = \Delta_i^{(e)}$$

$$\frac{\partial E}{\partial w_{ij}^{(e)}} = \Delta_i^{(e)} a_j^{(e-1)}$$

We now turn to the implementation in Mathematica which we can try to follow the steps explicitly.