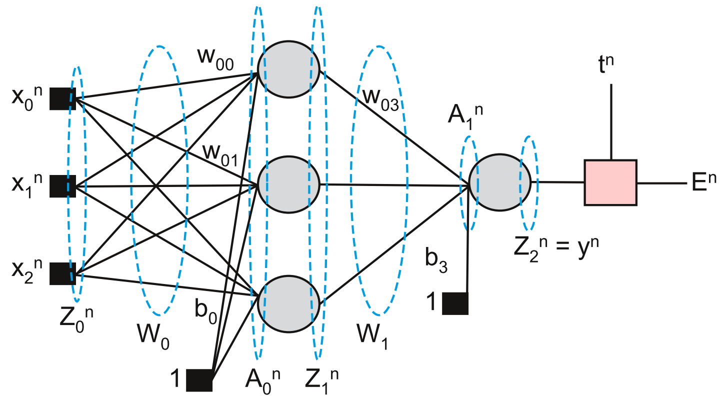

Figure 1: Neural Network example scheme

1 Forward

$a_0^n = w_{00} \cdot x_0^n + w_{10} \cdot x_1^n + w_{20} \cdot x_2^n + b_0$

$a_1^n = w_{01} \cdot x_0^n + w_{11} \cdot x_1^n + w_{21} \cdot x_2^n + b_1$

$a_2^n = w_{02} \cdot x_0^n + w_{12} \cdot x_1^n + w_{22} \cdot x_2^n + b_0$

$$\underset{Z_0[N \times 4]}{\begin{pmatrix} 1 & x_0^0 & x_1^1 & x_2^2 \\ 1 & x_0^0 & x_1^1 & x_2^2 \\ ... & & & \\ 1 & x_0^n & x_1^n & x_2^n \\ ... & & & \\ 1 & x_0^N & x_1^N & x_2^N \end{pmatrix}} \times \underset{W_0[4 \times 3]}{\begin{pmatrix} b_0 & b_1 & b_2 \\ w_{00} & w_{01} & w_{03} \\ w_{10} & w_{11} & w_{13} \\ w_{20} & w_{21} & w_{23} \end{pmatrix}} = \underset{A_0[N \times 3]}{\begin{pmatrix} a_0^0 & a_1^0 & a_2^0 \\ a_0^1 & a_1^1 & a_2^1 \\ ... & & \\ a_0^n & a_1^n & a_2^n \\ ... & & \\ a_0^N & a_1^N & a_2^N \end{pmatrix}}$$

$Z_1 = \sigma(A_0)$

$a_3^n = w_{03} \cdot z_0^n + w_{13} \cdot z_1^n + w_{23} \cdot z_2^n + b_3$

$$\underset{Z_1[N \times 4]}{\begin{pmatrix} 1 & z_0^0 & z_1^1 & z_2^2 \\ 1 & z_0^0 & z_1^1 & z_2^2 \\ ... & & & \\ 1 & z_0^n & z_1^n & z_2^n \\ ... & & & \\ 1 & z_0^N & z_1^N & z_2^N \end{pmatrix}} \times \underset{W_1[4 \times 1]}{\begin{pmatrix} b_3 \\ w_{03} \\ w_{13} \\ w_{23} \end{pmatrix}} = \underset{A_1[N \times 1]}{\begin{pmatrix} a_3^0 \\ a_3^1 \\ ... \\ a_3^n \\ ... \\ a_3^N \end{pmatrix}}$$

$Z_2 = \sigma(A_1)$

$E^n = \frac{1}{2}(y^n - t^n)^2 = \frac{1}{2}(z_2^n - t^n)^2$

$\sigma(x) = \frac{1}{1 + e^{-x}}$

---

2 Backpropagation

$$\frac{dE^n}{db_3} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{db_3} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot 1$$

$$\frac{dE^n}{dw_{03}} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dw_{03}} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot z_0^n$$

$$\frac{dE^n}{dw_{13}} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dw_{13}} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot z_0^n$$

$$\frac{dE^n}{dw_{23}} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dw_{23}} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot z_0^n$$

$$\delta_3^n = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n)$$

$$e^n = \frac{dE^n}{dz_3^n} = (z_3^n - t^n)$$

$$\underset{err[N \times 1]}{\begin{pmatrix} e^0 \\ e^1 \\ ... \\ e^n \\ ... \\ e^N \end{pmatrix}} \cdot \underset{[N \times 1]}{\begin{pmatrix} z_3^0 \\ z_3^1 \\ ... \\ z_3^n \\ ... \\ z_3^N \end{pmatrix}} \cdot \left(1 - \underset{[N \times 1]}{\begin{pmatrix} z_3^0 \\ z_3^1 \\ ... \\ z_3^n \\ ... \\ z_3^N \end{pmatrix}} \right) = \underset{\Delta_1[N \times 1]}{\begin{pmatrix} \delta_3^0 \\ \delta_3^1 \\ ... \\ \delta_3^n \\ ... \\ \delta_3^N \end{pmatrix}}$$

$$\underset{[4 \times N]}{\left(Z_1^T\right)} \times \underset{[N \times 1]}{\left(\Delta_1\right)} = \underset{[4 \times 1]}{\left(\tfrac{dE}{dW_1}\right)}$$

$$\frac{dE^n}{db_0} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dz_0^n} \cdot \frac{dz_0^n}{da_0^n} \cdot \frac{da_0^n}{db_0} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot w_{03} \cdot z_0^n \cdot (1 - z_0^n) \cdot 1$$

$$\frac{dE^n}{dw_{00}} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dz_0^n} \cdot \frac{dz_0^n}{da_0^n} \cdot \frac{da_0^n}{dw_{00}} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot w_{03} \cdot z_0^n \cdot (1 - z_0^n) \cdot x_0^n$$

...

$$\frac{dE^n}{dw_{11}} = \frac{dE^n}{dz_3^n} \cdot \frac{dz_3^n}{da_3^n} \cdot \frac{da_3^n}{dz_1^n} \cdot \frac{dz_1^n}{da_1^n} \cdot \frac{da_1^n}{dw_{11}} = (z_3^n - t^n) \cdot z_3^n \cdot (1 - z_3^n) \cdot w_{13} \cdot z_1^n \cdot (1 - z_1^n) \cdot x_1^n$$

$$\delta_0^n = \delta_3^n \cdot \frac{da_3^n}{dz_0^n} \cdot \frac{dz_0^n}{da_0^n} = \delta_3^n \cdot w_{03} \cdot z_0^n \cdot (1 - z_0^n)$$

$$\delta_1^n = \delta_3^n \cdot \frac{da_3^n}{dz_1^n} \cdot \frac{dz_1^n}{da_1^n} = \delta_3^n \cdot w_{13} \cdot z_1^n \cdot (1 - z_1^n)$$

$\delta_2^n = \delta_3^n \cdot \frac{da_3^n}{dz_2^n} \cdot \frac{dz_2^n}{da_2^n} = \delta_3^n \cdot w_{23} \cdot z_2^n \cdot (1 - z_2^n)$

...

$$\underset{\Delta_1[N \times 1]}{\begin{pmatrix} \delta_3^0 \\ \delta_3^1 \\ ... \\ \delta_3^n \\ ... \\ \delta_3^N \end{pmatrix}} \times \underset{[1 \times 3]}{\begin{pmatrix} w_{03} & w_{13} & w_{23} \end{pmatrix}} \cdot \underset{Z_0 nobias[N \times 3]}{\begin{pmatrix} z_0^0 & z_1^1 & z_2^2 \\ z_0^0 & z_1^1 & z_2^2 \\ ... \\ z_0^n & z_1^n & z_2^n \\ ... \\ z_0^N & z_1^N & z_2^N \end{pmatrix}} \cdot \left(1 - \underset{Z_0 nobias[N \times 3]}{\begin{pmatrix} z_0^0 & z_1^1 & z_2^2 \\ z_0^0 & z_1^1 & z_2^2 \\ ... \\ z_0^n & z_1^n & z_2^n \\ ... \\ z_0^N & z_1^N & z_2^N \end{pmatrix}} \right) = \underset{\Delta_0[N \times 3]}{\begin{pmatrix} \delta_0^0 & \delta_1^0 & \delta_2^0 \\ \delta_0^1 & \delta_1^1 & \delta_2^1 \\ ... \\ \delta_0^n & \delta_1^n & \delta_2^n \\ ... \\ \delta_0^N & \delta_1^N & \delta_2^N \end{pmatrix}}$$

$$\underset{[4 \times N]}{\begin{pmatrix} Z_0^T \end{pmatrix}} \times \underset{[N \times 3]}{\begin{pmatrix} \Delta_0 \end{pmatrix}} = \underset{[4 \times 3]}{\begin{pmatrix} \frac{dE}{dW_0} \end{pmatrix}}$$