Homework 7

CS1675
Professor Adriana Kovashka

Brian Knotten

November 15, 2018

Part I:

$x = [11]$, $y = [00]$, $\eta = 0.3$, $\delta_j = z_j(1 - z_j)\Sigma_k W_{kj}\delta_k$
Activation $= \frac{1}{1+e^{-a}}$, $a_j = \Sigma_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}$
All weights set to 0.05

Forward Activations:

$z_1 = \frac{1}{1+e^{-a}}$

$$\begin{aligned} -a &= x_0 w_{10}^{(1)} + x_1 w_{11}^{(1)} + x_2 w_{12}^{(1)} \\ &= 1 \cdot 0.05 + 1 \cdot 0.05 + 1 \cdot 0.05 \\ &= 0.15 \end{aligned}$$

$z_1 = \frac{1}{1+e^{-0.15}} = 0.5374$

$z_2 = \frac{1}{1+e^{-a}}$

$$\begin{aligned} -a &= x_0 w_{20}^{(1)} + x_1 w_{21}^{(1)} + x_2 w_{22}^{(1)} \\ &= 1 \cdot 0.05 + 1 \cdot 0.05 + 1 \cdot 0.05 \\ &= 0.15 \end{aligned}$$

$z_2 = \frac{1}{1+e^{-0.15}} = 0.5374$

$y_1 = \frac{1}{1+e^{-a}}$

$$\begin{aligned} -a &= z_0 w_{10}^{(2)} + z_1 w_{11}^{(2)} + z_2 w_{12}^{(2)} \\ &= 1 \cdot 0.05 + 0.5374 \cdot 0.05 + 0.5374 \cdot 0.05 \\ &= 0.1037 \end{aligned}$$

$y_1 = \frac{1}{1+e^{-0.1037}} = 0.5259$

$y_2 = \frac{1}{1+e^{-a}}$

$$\begin{aligned} -a &= z_0 w_{20}^{(2)} + z_1 w_{21}^{(2)} + z_2 w_{22}^{(2)} \\ &= 1 \cdot 0.05 + 0.5374 \cdot 0.05 + 0.5374 \cdot 0.05 \\ &= 0.1037 \end{aligned}$$

$y_2 = \frac{1}{1+e^{-0.1037}} = 0.5259$

Backpropogation:

$\delta_k = y_k(1-y_k)(y_k - t_k)$, t_k = desired y_k output = 0, $y_k = y_1 = y_2 = 0.5259$
$\frac{\partial E}{\partial w_{kj}^{(2)}} = \delta_k z_j$, $w^{(\tau+1)} = w^{(\tau)} - \eta \nabla E(w^{(\tau)})$

$\delta_{y1} = 0.5259 \cdot (1 - 0.5259) \cdot (0.5259 - 0) = 0.1311$
$\delta_{y2} = 0.5259 \cdot (1 - 0.5259) \cdot (0.5259 - 0) = 0.1311$
$z_0 = 1, z_1 = z_2 = 0.5374, \eta = 0.3, w^{(\tau)} = 0.05 \ \forall w_{kj}$

$w_{10}^{(2)} = w_{10}^{(2)} - 0.3 \cdot \delta_{y1} \cdot z_0 = 0.05 - 0.3 \cdot 0.1311 \cdot 1 = 0.01067$
$w_{11}^{(2)} = w_{11}^{(2)} - 0.3 \cdot \delta_{y1} \cdot z_1 = 0.05 - 0.3 \cdot 0.1311 \cdot 0.5374 = 0.02886$
$w_{12}^{(2)} = w_{12}^{(2)} - 0.3 \cdot \delta_{y1} \cdot z_2 = 0.05 - 0.3 \cdot 0.1311 \cdot 0.5374 = 0.02886$

$w_{20}^{(2)} = w_{20}^{(2)} - 0.3 \cdot \delta_{y2} \cdot z_0 = 0.05 - 0.3 \cdot 0.1311 \cdot 1 = 0.01067$
$w_{21}^{(2)} = w_{21}^{(2)} - 0.3 \cdot \delta_{y2} \cdot z_1 = 0.05 - 0.3 \cdot 0.1311 \cdot 0.5374 = 0.02886$
$w_{22}^{(2)} = w_{22}^{(2)} - 0.3 \cdot \delta_{y2} \cdot z_2 = 0.05 - 0.3 \cdot 0.1311 \cdot 0.5374 = 0.02886$

$\delta_j = z_j(1-z_j)\Sigma_k w_{kj}\delta_k$, $z_j = z_j$ from activation (0.5374), w_{kj} = weight from before update (0.05), $\frac{\partial E}{\partial w_{ji}^{(1)}} = \delta_j x_i$

$$\begin{aligned}\delta_{z1} &= z_1 \cdot (1 - z_1) \cdot (w_{11}^{(2)} \cdot \delta_{y1} + w_{21}^{(2)} \cdot \delta_{y2}) \\ &= 0.5374 \cdot (1 - 0.5374) \cdot (0.05 \cdot 0.1311 + 0.05 \cdot 0.1311) \\ &= 0.003259\end{aligned}$$

$$\begin{aligned}\delta_{z2} &= z_2 \cdot (1 - z_2) \cdot (w_{12}^{(2)} \cdot \delta_{y1} + w_{22}^{(2)} \cdot \delta_{y2}) \\ &= 0.5374 \cdot (1 - 0.5374) \cdot (0.05 \cdot 0.1311 + 0.05 \cdot 0.1311) \\ &= 0.003259\end{aligned}$$

$w_{10}^{(1)} = w_{10}^{(1)} - 0.3 \cdot \delta_{z1} \cdot x_0 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$
$w_{11}^{(1)} = w_{11}^{(1)} - 0.3 \cdot \delta_{z1} \cdot x_1 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$
$w_{12}^{(1)} = w_{12}^{(1)} - 0.3 \cdot \delta_{z1} \cdot x_2 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$

$w_{20}^{(1)} = w_{20}^{(1)} - 0.3 \cdot \delta_{z2} \cdot x_0 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$
$w_{21}^{(1)} = w_{21}^{(1)} - 0.3 \cdot \delta_{z2} \cdot x_1 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$
$w_{22}^{(1)} = w_{22}^{(1)} - 0.3 \cdot \delta_{z2} \cdot x_2 = 0.05 - 0.3 \cdot 0.003259 \cdot 1 = 0.04902$