Homework 5

Discussion of Question 1:

$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$:

$$N_1 = \sigma(w_{11} \times x_1 + w_{21} \times x_2) = \sigma(0.2) = \frac{1}{1+e^{-0.2}} = 0.55$$

$$N_2 = \sigma(w_{12} \times x_1 + w_{22} \times x_2) = \sigma(0.95) = \frac{1}{1+e^{-0.95}} = 0.72$$

$$y_1 = \sigma(w_{13} \times N_1 + w_{23} \times N_2) = \sigma(0.652) = \frac{1}{1+e^{-0.652}} = 0.657$$

$$y_2 = \sigma(w_{14} \times N_1 + w_{24} \times N_2) = \sigma(0.491) = \frac{1}{1+e^{-0.491}} = 0.62$$

$$w_{13}^{(1)} = w_{13}^{(0)} + \eta e_1 N_1 y_1 (1-y_1) = 0.4 + 0.1 \cdot (1-0.657) \cdot 0.55 \cdot 0.657 \cdot (1-0.657) = 0.404$$

$$w_{23}^{(1)} = w_{23}^{(0)} + \eta e_1 N_1 y_1 (1-y_1) = 0.6 + 0.1 \cdot (1-0.657) \cdot 0.55 \cdot 0.657 \cdot (1-0.657) = 0.604$$

$$w_{14}^{(1)} = w_{14}^{(0)} + \eta e_2 N_2 y_2 (1-y_2) = 0.5 + 0.1 \cdot (1-0.62) \cdot 0.72 \cdot 0.62 \cdot (1-0.62) = 0.506$$

$$w_{24}^{(1)} = w_{24}^{(0)} + \eta e_2 N_2 y_2 (1-y_2) = 0.3 + 0.1 \cdot (1-0.62) \cdot 0.72 \cdot 0.62 \cdot (1-0.62) = 0.306$$

$$w_{11}^{(1)} = w_{11}^{(0)} + \eta x_1 N_1 (1-N_1)[e_1 y_1 (1-y_1) w_{13}^{(0)} + e_2 y_2 (1-y_2) w_{14}^{(0)}]$$
$$= 0.1 + 0.1 \cdot 1 \cdot 0.55 \cdot (1-0.55)[(1-0.657) \cdot 0.657 \cdot (1-0.657) \cdot 0.4 + (0-0.62) \cdot 0.62 \cdot (1-0.62) \cdot 0.5]$$
$$= 0.0990$$

$$w_{21}^{(1)} = w_{21}^{(0)} + \eta x_1 N_1 (1-N_1)[e_1 y_1 (1-y_1) w_{13}^{(0)} + e_2 y_2 (1-y_2) w_{14}^{(0)}]$$
$$= 0.1 + 0.1 \cdot 1 \cdot 0.55 \cdot (1-0.55)[(1-0.657) \cdot 0.657 \cdot (1-0.657) \cdot 0.4 + (0-0.62) \cdot 0.62 \cdot (1-0.62) \cdot 0.5]$$
$$= 0.0990$$

$$w_{12}^{(1)} = w_{12}^{(0)} + \eta x_2 N_2 (1-N_2)[e_1 y_1 (1-y_1) w_{23}^{(0)} + e_2 y_2 (1-y_2) w_{24}^{(0)}]$$
$$= 0.25 + 0.1 \cdot 1 \cdot 0.72 \cdot (1-0.72)[(1-0.657) \cdot 0.657 \cdot (1-0.657) \cdot 0.6 + (0-0.62) \cdot 0.62 \cdot (1-0.62) \cdot 0.3]$$
$$= 0.2501$$

$$w_{22}^{(1)} = w_{22}^{(0)} + \eta x_2 N_2 (1-N_2)[e_1 y_1 (1-y_1) w_{23}^{(0)} + e_2 y_2 (1-y_2) w_{24}^{(0)}]$$
$$= 0.7 + 0.1 \cdot 1 \cdot 0.72 \cdot (1-0.72)[(1-0.657) \cdot 0.657 \cdot (1-0.657) \cdot 0.6 + (0-0.62) \cdot 0.62 \cdot (1-0.62) \cdot 0.3]$$
$$= 0.7001$$

Discussion of Question 2:

$$E(n)=\frac{1}{2}\sum_{n=1}^{N}e_n^2=\sum_{n=1}^{N}(d_n-y_n)^2=\sum_{n=1}^{N}(d_n-\phi_l(v_l(n)))^2$$

$$\phi(v)=\tanh(v)=\frac{e^v-e^{-v}}{e^v+e^{-v}}$$

$$\phi^{'}(v)=\frac{4}{(e^v+e^{-v})^2}=(1\text{-}\phi(v))(1+\phi(v))$$

$$\frac{\partial E(n)}{\partial w_{lj}}=\frac{\partial E(n)}{\partial e_l(n)}\frac{\partial e_l(n)}{\partial y_l(n)}\frac{\partial y_l(n)}{\partial v_l(n)}\frac{\partial v_l(n)}{\partial w_{lj}}$$

$$=[e_l][-1][(1-y_l(n))(1+y_l(n))][y_{jl}(n)]$$

$$w_{il}(n+1)=w_{il}(n)+\eta e_l[(1-y_l(n))(1+y_l(n))][y_i]$$

Discussion of Question 3:

$$E(n)=\frac{1}{2}\sum_{n=1}^{N}e_n^2=\sum_{n=1}^{N}(d_n-y_n)^2=\sum_{n=1}^{N}(d_n-\phi_l(v_l(n)))^2$$

$$\phi(v_n)=\frac{e^{v_n}}{\sum_{i=1}^{O}e^{v_i}}$$

$$\frac{\partial\phi(v_n)}{\partial v_j}=\begin{cases}\dfrac{e^{v_n}(\sum_{i=1}^{O}e^{v_i})-e^{v_n}\cdot e^{v_n}}{(\sum_{i=1}^{O}e^{v_i})^2}, j=n\\[2ex] \dfrac{0\cdot(\sum_{i=1}^{O}e^{v_i})-e^{v_n}\cdot e^{v_j}}{(\sum_{i=1}^{O}e^{v_i})^2}, j\neq n\end{cases}$$

$$\therefore\frac{\partial\phi(v_n)}{\partial v_j}=\begin{cases}\phi(v_n)(1-\phi(v_n)), j=n\\ -\phi(v_n)\phi(v_j), j\neq n\end{cases}$$

$$\frac{\partial E(n)}{\partial w_{lj}}=\frac{\partial E(n)}{\partial e_l(n)}\frac{\partial e_l(n)}{\partial y_l(n)}\frac{\partial y_l(n)}{\partial v_l(n)}\frac{\partial v_l(n)}{\partial w_{lj}}$$

$$=[d_l-y_l][-1][y_l(n)(1-y_l(n))][y_{jl}(n)]$$

$$w_{jk}(n+1) = w_{jk}(n) + \eta e_k \left[\frac{e^{v_k} (\sum_{i=1}^{O} e^{v_i}) - e^{v_k} \cdot e^{v_k}}{(\sum_{i=1}^{O} e^{v_i})^2} \right] [y_j]$$

$$= w_{jk}(n) + \eta (d_k - y_k) [y_k(n)(1 - y_k(n))] [y_j]$$