1)

a) $\boxed{p=2}$ as the input has i1 and i2

b) $h_1 = \sigma(w_1 i_1 + w_2 i_2 + b_1)$

$= \sigma(0.15 \cdot 0.05 + 0.20 \cdot 0.10 + 0.35)$

$= \sigma(0.3775)$

$= 0.5933$

$h_2 = \sigma(w_3 i_1 + w_4 i_2 + b_1)$

$= \sigma(0.25 \cdot 0.05 + 0.30 \cdot 0.10 + 0.35)$

$= \sigma(0.3925)$

$= 0.5969$

$\hat{o}_1 = \sigma(w_5 h_1 + w_6 h_2 + b_2)$

$= \sigma(0.40 \cdot 0.5933 + 0.45 + 0.5969 + 0.60)$

$= \sigma(1.1059)$

$= 0.7514$

$\hat{o}_2 = \sigma(w_7 h_1 + w_8 h_2 + b_2)$

$= \sigma(0.50 \cdot 0.5933 + 0.55 \cdot 0.5969 + 0.60)$

$= \sigma(1.2249)$

$= 0.7729$

$\boxed{\hat{o}_1 = 0.7514 \quad \hat{o}_2 = 0.7729}$

c) $o_1 = 0.01 \quad o_2 = 0.99$

$L = \frac{1}{2} \sum_{i=1}^{2} (\hat{o}_i - o_i)^2$

$= \frac{1}{2}\left((0.7514 - 0.01)^2 + (0.7729 - 0.09)^2\right)$

$= 0.2984$

$\boxed{L = 0.2984}$

d)

$$W^1=\begin{pmatrix} w_1 & w_2 \\ w_3 & w_4 \end{pmatrix} \quad W^2=\begin{pmatrix} w_5 & w_6 \\ w_7 & w_8 \end{pmatrix}$$

$$b^1=\begin{pmatrix} b_1 \\ b_1 \end{pmatrix} \quad b^2=\begin{pmatrix} b_2 \\ b_2 \end{pmatrix} \quad y=\begin{pmatrix} o_1 \\ o_2 \end{pmatrix}$$

$$a^0=\begin{pmatrix} i_1 \\ i_2 \end{pmatrix}$$

$$z^1=W^1a^0+b^1$$
$$=\begin{pmatrix} w_1 & w_2 \\ w_3 & w_4 \end{pmatrix}\begin{pmatrix} i_1 \\ i_2 \end{pmatrix}+\begin{pmatrix} b_1 \\ b_1 \end{pmatrix}$$
$$=\begin{pmatrix} w_1i_1+w_2i_2+b_1 \\ w_3i_1+w_4i_4+b_1 \end{pmatrix}=\begin{pmatrix} h_1^{in} \\ h_2^{in} \end{pmatrix}=\begin{pmatrix} 0.3775 \\ 0.3925 \end{pmatrix}$$

$$a^1=\sigma(z^1)$$
$$=\begin{pmatrix} \sigma(h_1^{in}) \\ \sigma(h_2^{in}) \end{pmatrix}=\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=\begin{pmatrix} 0.5933 \\ 0.5969 \end{pmatrix}$$

$$z^2=W^2a^1+b^2$$
$$=\begin{pmatrix} w_5 & w_6 \\ w_7 & w_8 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}+\begin{pmatrix} b_2 \\ b_2 \end{pmatrix}$$
$$=\begin{pmatrix} w_5h_1+w_6h_2+b_2 \\ w_7h_1+w_8h_2+b_2 \end{pmatrix}=\begin{pmatrix} o_1^{in} \\ o_2^{in} \end{pmatrix}=\begin{pmatrix} 1.1059 \\ 1.2249 \end{pmatrix}$$

$$a^2=\sigma(z^2)$$
$$=\begin{pmatrix} \sigma(o_1^{in}) \\ \sigma(o_2^{in}) \end{pmatrix}=\begin{pmatrix} \hat{o}_1 \\ \hat{o}_2 \end{pmatrix}=\begin{pmatrix} 0.7514 \\ 0.7729 \end{pmatrix}$$

$$\delta^2=(a^2-y)\odot\sigma'(z^2)$$
$$=\begin{pmatrix} \hat{o}_1-o_1 \\ \hat{o}_2-o_2 \end{pmatrix}\odot\begin{pmatrix} \sigma(o_1^{in})(1-\sigma(o_1^{in})) \\ \sigma(o_2^{in})(1-\sigma(o_2^{in})) \end{pmatrix}$$
$$=\begin{pmatrix} (\hat{o}_1-o_1)\sigma(o_1^{in})(1-\sigma(o_1^{in})) \\ (\hat{o}_2-o_2)\sigma(o_2^{in})(1-\sigma(o_2^{in})) \end{pmatrix}$$
$$=\begin{pmatrix} (0.7514-0.01)\,\sigma(1.1059)(1-\sigma(1.1059)) \\ (0.7229-0.99)\,\sigma(1.2249)(1-\sigma(1.2249)) \end{pmatrix}=\begin{pmatrix} 0.1385 \\ -0.0381 \end{pmatrix}=\begin{pmatrix} \delta_1^2 \\ \delta_2^2 \end{pmatrix}$$

$$\nabla f(W^2)=\delta^2(a^1)^T$$
$$=\begin{pmatrix} \delta_1^2 \\ \delta_2^2 \end{pmatrix}(h_1\ h_2)$$
$$=\begin{pmatrix} 0.1385 \\ -0.0381 \end{pmatrix}(0.5933\ \ 0.5969)=\begin{pmatrix} 0.0822 & 0.0827 \\ -0.0226 & -0.0227 \end{pmatrix}$$

$$\delta^1=(W^2)^T\delta^2\odot\sigma'(z^1)$$
$$=\begin{pmatrix} w_5 & w_7 \\ w_6 & w_8 \end{pmatrix}\begin{pmatrix} \delta_1^2 \\ \delta_2^2 \end{pmatrix}\odot\begin{pmatrix} \sigma(h_1^{in})(1-\sigma(h_1^{in})) \\ \sigma(h_2^{in})(1-\sigma(h_2^{in})) \end{pmatrix}$$
$$=\begin{pmatrix} (w_5\delta_1^2+w_7\delta_2^2)\,\sigma(h_1^{in})(1-\sigma(h_1^{in})) \\ (w_6\delta_1^2+w_8\delta_2^2)\,\sigma(h_2^{in})(1-\sigma(h_2^{in})) \end{pmatrix}=\begin{pmatrix} \delta_1^1 \\ \delta_2^1 \end{pmatrix}$$
$$=\begin{pmatrix} (0.40\cdot 0.1385+0.50\cdot -0.0381)\,\sigma(0.3775)(1-\sigma(0.3775)) \\ (0.45\cdot 0.1385+0.55\cdot -0.0381)\,\sigma(0.3925)(1-\sigma(0.3925)) \end{pmatrix}=\begin{pmatrix} 0.0088 \\ 0.001 \end{pmatrix}$$

$$\nabla f(W^1)=\delta^1(a^0)^T\odot\sigma'(z^1)$$
$$=\begin{pmatrix} \delta_1^1 \\ \delta_2^1 \end{pmatrix}(i_1\ i_2)$$
$$=\begin{pmatrix} 0.0088 \\ 0.001 \end{pmatrix}(0.05\ \ 0.10)=\begin{pmatrix} 4.3857\times10^{-4} & 8.7714\times10^{-4} \\ 4.9771\times10^{-4} & 9.9543\times10^{-4} \end{pmatrix}$$

$$\boxed{\begin{array}{ll} \frac{\partial L}{\partial w_1}=4.3857\times10^{-4} & \frac{\partial L}{\partial w_2}=8.7714\times10^{-4} \\ \frac{\partial L}{\partial w_3}=4.9771\times10^{-4} & \frac{\partial L}{\partial w_4}=9.9543\times10^{-4} \\ \frac{\partial L}{\partial w_5}=0.0822 & \frac{\partial L}{\partial w_6}=0.0827 \\ \frac{\partial L}{\partial w_7}=-0.0226 & \frac{\partial L}{\partial w_8}=-0.0227 \end{array}}$$

e) $\boxed{w_k^{(t+1)}=w_k^{(t)}-\tau\cdot\frac{\partial L}{\partial w_k}}$

where we are on iteration $t+1$ and updating weight k,

$\frac{\partial L}{\partial w_k}$ was calculated in part (d)

and assume τ is given