1.1 $E = MSE = \frac{1}{2}(\text{target} - o)^2$

Partial derivative for wi $= \frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial o} * \frac{\partial o}{\partial w_i}$

$\frac{\partial o}{\partial w_i} = (x_i + x_i^2)$

w_i is used as $w_i(x_i + x_i^2)$
So derivative is $(x_i + x_i^2)$

$\frac{\partial E}{\partial o} = \frac{\partial \frac{1}{2}(\text{target} - o)^2}{\partial o} = \frac{(\text{target} - o)\,\partial(\text{target} - o)}{\partial o}$

taking $\frac{\partial(\text{target} - o)}{\partial o}$

$= \frac{\partial \text{target}}{\partial o} - \frac{\partial o}{\partial o} \rightarrow \frac{\partial \text{target}}{\partial o} = 0 \quad \frac{\partial o}{\partial o} = 1$

therefore $\frac{\partial E}{\partial w_i} = (\text{target} - o)(-1) = -(\text{target} - o)$

putting together $\frac{\partial E}{d w_i} = -(\text{target} - o) * (x_i + x_i^2)$

$w_{n\,new} = w_n - \eta\left(\frac{\partial E}{d w_i}\right) \Rightarrow w_n - \eta\left[(\text{target} + o)(x_i + x_i^2)\right]$

So weight update $w_{n\,new} = w_n - \eta\left[(\text{target} + o)(x_i + x_i^2)\right]$

1.2 a) $N3 = o((x_1 * w_{31}) + (x_2 * w_{32}))$

N3 = neuron 3
N4 = neuron 4
N5 = neuron 5

$N4 = o((x_1 * w_{41}) + (x_2 * w_{42}))$

$N5 = o((N3 * w_{53}) + (N4 * w_{54}))$

$y_5 = \left[o((x_1 * w_{31}) + (x_2 * w_{32})) * w_{53} + o((x_1 * w_{41}) + (x_2 * w_{42})) * w_{54} \right]$

I used o to indicate output from that neuron

b) output from input layer = X

output from hidden layer = $(X * W^{(1)})$

output from output layer = $(W^{(2)} * (X * W^{(1)}))$

So output of neural network is $(W^{(2)} * (X * W^{(1)}))$

c) $h_s(x) = \frac{1}{1+e^{-x}}$ $h_t(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

we take $h_s(x) \approx h_t(x)$

- as x goes to ∞

$$\frac{1}{1+e^{-x}} \approx \frac{e^x - e^{-x}}{e^x + e^{-x}} \rightarrow \frac{1}{1+0} \approx \frac{\infty - 0}{\infty + 0} \Rightarrow 1 \approx 1$$

- or, as x goes to $-\infty$

$$\frac{1}{1+e^{-x}} \approx \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \Rightarrow \frac{1}{1+\infty} \approx \frac{0-\infty}{0+\infty} \Rightarrow 0 \approx 0$$

So we can prove that neural nets can created using above to functions can generate the same function by taking limit as x goes to ∞ or $-\infty$