Question 1

(a) We need to show $J_{MSE}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\left(\hat{y}^{(n)} - y^{(n)}\right)^2$ is not convex.

For simplicity:

$$f(x) = (y - \hat{y})^2 \quad \text{and} \quad \hat{y} = \frac{1}{1+e^{-\theta x}} \longrightarrow \text{sigmoid function}$$

$$g(x) = \frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \theta}$$

derivative

$$= -2(y - \hat{y})\,\hat{y}(1-\hat{y})x$$

$$g(x) = \frac{\partial f}{\partial \theta} = -2\left[y\hat{y} - y\hat{y}^2 - \hat{y}^2 + \hat{y}^3\right]x$$

$$\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial \theta}\right) = \frac{\partial g}{\partial \theta} = \frac{\partial g}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \theta}$$

$$= -2\left[y - 2y\hat{y} - 2\hat{y} + 3\hat{y}^2\right]x \cdot x\,\hat{y}(1-\hat{y})$$

always between $[0, 1/4]$ since $\hat{y}:[0,1]$

$$H(\hat{y}) = -2\left[3\hat{y}^2 - 2\hat{y}(y+1) + y\right]x^2 \longrightarrow \text{always positive}$$

From above it can be seen that **ŷ * (1 - ŷ)** lies between [0, 1]

Hence we have to check that if H(ŷ) is positive for all values of "x" or not

When $y = 0$

we have $H(\hat{y}) = -2[3\hat{y}^2 - 2\hat{y}(y+1) + y]$

$$H(\hat{y}) = -2[3\hat{y}^2 - 2\hat{y}]$$

$$= -2\left[3\hat{y}\left(\hat{y} - \frac{2}{3}\right)\right]$$

When $y = 1$

$$H(\hat{y}) = -2[3\hat{y}^2 - 4\hat{y} + 1]$$

by factorizing we get

$$= -2\left[3\left(\hat{y} - \frac{1}{3}\right)(\hat{y} - 1)\right]$$

For y = 0, it is clear from the equation that when ŷ lies in the range **[0, 2/3]** the function **H(ŷ) ≥ 0 and** when ŷ lies between **[2/3, 1]** the function **H(ŷ) ≤ 0.** This shows the function is not convex.

For y = 1, it is clear from the equation that when ŷ lies in the range **[0, 1/3]** the function **H(ŷ) ≤ 0 and** when ŷ lies between **[1/3, 1]** the function **H(ŷ) ≥ 0.** This also shows the function is not convex.

Hence the mean square error cost function is NOT convex.

(b) We need to show that log-loss cost function is a CONVEX function

We proceed by simplifying the expression a little

$$-f(x) = y\log(\hat{y}) + (1-y)\log(1-\hat{y})$$

$$= y\log\left(\frac{1}{1+e^{-\theta x}}\right) + (1-y)\log\left(1-\frac{1}{1+e^{-\theta x}}\right)$$

$$= y\log\left(\frac{e^{\theta x}}{1+e^{\theta x}}\right) + (1-y)\log\left(\frac{1}{1+e^{\theta x}}\right)$$

$$= y\log(e^{\theta x}) - y\log(1+e^{\theta x})$$

$$+ \cancel{(1-y)\log(1)}^{\to 0} - (1-y)\log(1+e^{\theta x})$$

$$= y(\theta x) - \cancel{y\log(1+e^{\theta x})}$$

$$-\log(1+e^{\theta x}) + \cancel{y\log(1+e^{\theta x})}$$

$$-f(x) = xy\theta - \log(1+e^{\theta x})$$

$$f(x) = \log(1+e^{\theta x}) - xy\theta$$

$$\frac{\partial f}{\partial \theta} = \frac{1}{1+e^{\theta x}}\ x\cdot e^{\theta x} - xy$$

$$= \frac{x}{1+e^{-\theta x}} - xy$$

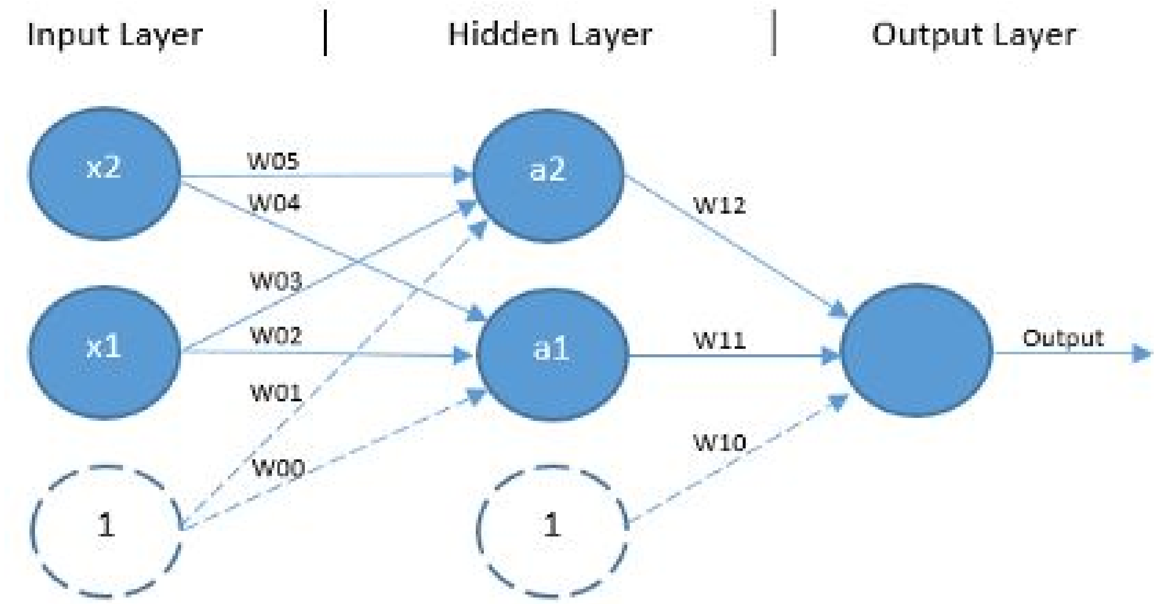
$$\frac{\partial^2 f}{\partial \theta^2} = x\cdot\frac{(-1)}{(1+e^{-\theta x})^2}(e^{-\theta x})(-x)$$

$$= \frac{x^2 e^{-\theta x}}{(1+e^{-\theta x})^2} \geq 0 \quad \forall x$$

Range of e^(x) = (0, ∞),
So is the final term. It is always positive. Hence it is a convex function.

Question 2:

(a)



(b) In order to show that this is the correct implementation of XOR; we follow the steps below

We can write

XOR(X1, X2) = NOR(NOR(X1, X2), AND(X1, X2))

= NOR (a1, a2)

(X1, X2)	a1=NOR, a2=AND	NOR(a1, a2)
0, 0	1, 0	0
0, 1	0, 0	1
0, 1	0, 0	1
1, 1	0, 1	0

As we can see NOR(a1, a2) in fact is same as XOR(x1, x2) as given in the truth table. Hence our implementation is the correct XOR implementation.