Deep Learning: HW2

1. Explain why "non-linearity" is required property for any activation function deployed in a deep neural network.
2. Derive $\partial L/\partial W_{x2,h1}$ using backward propagation