MLAI Week 7 Exercise: Neural Networks

Note: An indicative mark is in front of each question. The total mark is 12. You may mark your own work when we release the solutions.

[2] 1. Using the definitions for **o** and **h** on slide 10 of Lecture 7 to show that if the activation function is linear such that $g(a) = a$, then the one-hidden-layer on that slide encodes a linear relationship between the input **x** and output **o**. Include all steps.

[1] 2. In Slide 38: we change the 3×3 kernel to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. What will be the 3×3 convolved features? What features can this kernel detect?

[3] 3. We have a $512 \times 512 \times 3$ colour image. We apply 100 5×5 filters with stride 7, and pad 2 to obtain a convolution output. What is the output volume size? How many parameters are needed for such a layer?

[6] 4. For the AlexNet depicted in Slide 35 of Lecture 6, there are about 60 million learnable parameters. With the help of the illustration `https://static.packt-cdn.com/products/9781789956177/graphics/assets/ec08175c-5282-4be2-b6e7-6f2d99272166.png`, compute the exact number of learnable parameters in AlexNet, showing the steps.