

STAT448 Assignment 3

Due – 11.59 pm 1st June 2025, submitted directly through LEARN. You can work in pairs if you choose. Please show all your work.

Q1. A. (5 marks) The following matrices describing a neural network were uncovered by scientists.

The weights for the hidden layer are given in the matrix:

$$W^{[1]} = \begin{bmatrix} 1 & 0 \\ 2 & 6 \end{bmatrix}$$

The bias for the hidden layer is given in the vector:

$$b^{[1]} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

The weights for the output layer are given in the vector:

$$W^{[2]} = [-1 \quad 7]$$

The bias for the output layer is:

$$b^{[2]} = [5]$$

The input, X , is given in the vector:

$$X = \begin{bmatrix} 5 \\ -7 \end{bmatrix}$$

The activation function for all the units in the hidden and output layers is ReLu:

$$g(z) = \max(0, z)$$

Draw the diagram of this network and calculate the output.

B. (5 marks) The input of a convolutional neural network is a tiny image, I . Its size is 5 x 5 pixels. The pixel values are given below.

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

For the first convolutional layer, the network engineers decided to use three filters of size 3 x 3 with stride 1 and no padding. After the engineers trained the network for 7 nights and 7 days non-stop, the values of the filters emerged as the following:

$$F1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F3 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Unfortunately, the power went down. Help the engineers by applying the filters F1, F2 and F3 to the image *I*, above. Then apply the activation function (ReLU). Write down the resulting three feature maps. Draw a diagram showing the link between the input and the bottom right pixel of the output feature map through this convolution.

Q2. (20 marks) There are datasets that are a bit more challenging than MNIST, such as Fashion-MNIST or CIFAR10. They are available in Keras directly. See the documentation in Keras on how to load them; the function is similar to MNIST.

Using code from the recent labs (on training artificial neural networks in Keras), train several networks with different settings on this dataset. Overall, try to train up to 10 different configurations. You might try Optuna. Write a short report (up to 5 pages max.) describing what effect different settings had on test accuracy and the training process overall. For example, try different learning rates, optimisers, batch sizes and/or architectures.

Please have captions for all figures and tables and make sure that if you include a figure or a table, it's referred to somewhere in the text. Submit this report as a single PDF file. Please have a structure in the report with headings for the sections. All figures need to have a legend and/or be clearly labelled.

Any references should use APA style for both in-text citations and the reference list.