

# STAT448 Big Data

## Assignment 3

**Due - 2d of June (end of day) submitted directly through Learn. Assignment can be done in pairs. Please show your work.**

**Q1. A. (4 marks)** The following matrices describing a neural network were uncovered by scientists.

The weights for the hidden layer are given in the matrix

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

The bias for the hidden layer is given in the vector

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

The weights for the output layer are given in the vector

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

The biases for the output layer are

$$b^{[2]} = \begin{bmatrix} -0.5 \\ 0.75 \end{bmatrix}$$

The input  $X$  is given in the vector

$$X = \begin{bmatrix} 1.25 \\ -0.5 \end{bmatrix}$$

The activation function for all units in the hidden and output layers is Relu:  $g(z) = \max(0, z)$

Determine dimensions of the output, draw the diagram of this network and calculate the output.

**B. (4 marks)** The input of a convolutional neural network is a tiny image  $I$ . Its size is 5\*5 pixels. The pixel values are given below.

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

For the first convolutional layer, the network engineers decided to use three filters of size 3\*3 with the 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 following:

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

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

$$F3 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Unfortunately, the power went down. Help the engineers by applying the filters  $F1, F2$  and  $F3$  to the image  $I$  above. 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. (12 marks)** There are datasets that are a bit more challenging than MNIST, such as Fashion-MNIST. It is available in Keras directly. See documentation in Keras how to load it, the function is similar to MNIST.

Using the code of the recent labs (on training artificial neural networks in Keras), train several networks with different settings on this dataset. Write a short report (up to 5 pages max) describing what effect different settings had on test accuracy and training process overall.

For example, try different learning rates, optimisers, batch sizes and architectures.

Please have captions for all figures and tables and make sure that if you included a figure or a table, it's described somewhere in the text.

Submit this report as a single pdf file. Note that having structure in the report will be very beneficial.