Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
First Semester 2022-2023

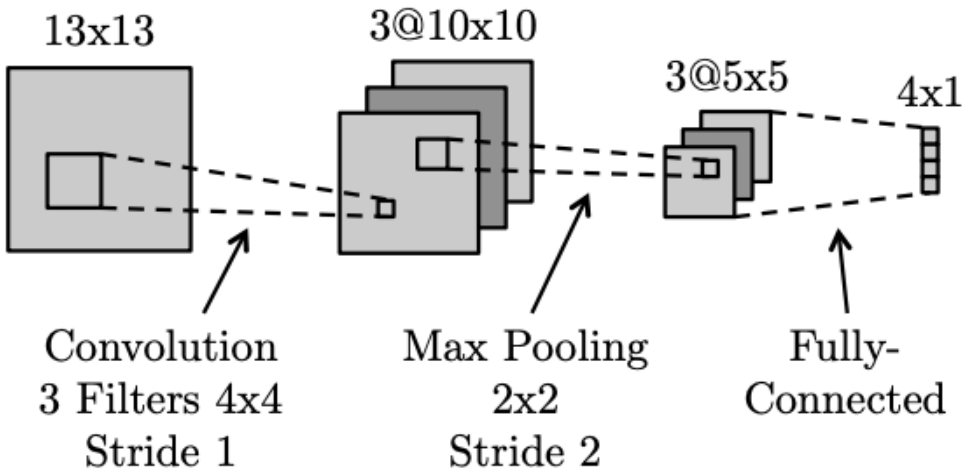
Mid-Semester Test
(EC-2 Regular)

Course No.	: SS ZG529
Course Title	: Deep Learning
Nature of Exam	: Open Book
Weightage	: 30%
Duration	: 2 Hours
Date of Exam	: 23/09/2022

No. of Pages = 2
No. of Questions = 4

Note to Students:
1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

Q1.
Below is a diagram of a small convolutional neural network that converts a 13x13 image into 4 output values. The network has the following layers/operations from input to output: convolution with 3 filters, max pooling, ReLu, and finally a fully-connected layer. For this network we will not be using any bias/offset parameters (b). Please answer the following questions about this network.

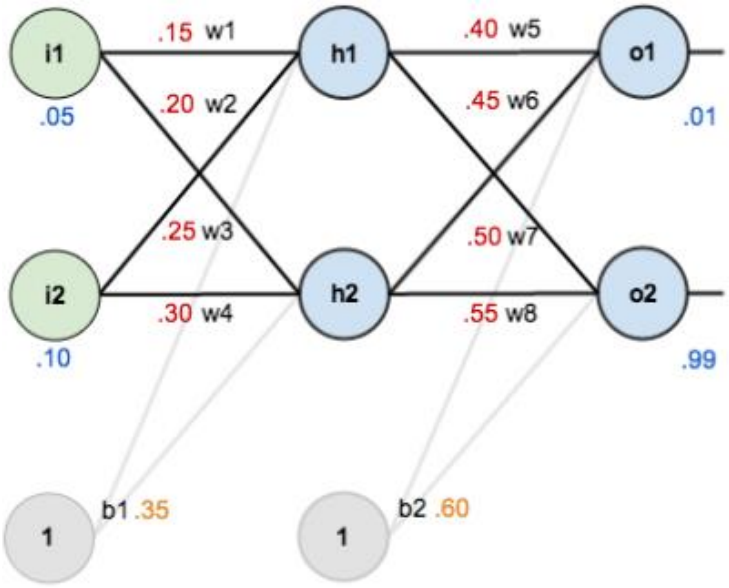


a) How many weights in the convolutional layer do we need to learn? [2]
b) How many ReLu operations are performed on the forward pass? [2]
c) How many weights do we need to learn for the entire network? [2]
d) Can a fully-connected neural network with the same size layers as the above network can represent any classifier that the above convolutional network can represent? [2]
e) What are the disadvantages of a fully-connected neural network compared to a convolutional neural network with the same size layers? [2]

Q2.
Train a perceptron using perceptron convergence algorithm to perform NOR logic operation. Choose initial weights such that the network goes through at least 1 weight update. (use a bias term) [8]

Q3.

Consider the network below:



(i1-h2=w3=.25, i2-h1=w2=.20, h2-o1=.45 and h1-o2=.50)

a) The network uses sigmoid activation function. Compute the total error that has to be back propagated through the network. [5]

b) Find the output at o1 and o2 if the activation function at the output layer is changed to Relu? [1]

c) Find the output of the network if a softmax layer is introduced after o1, o2 layer? [1]

Q4.

The image C was obtained by convolving the image I with a 2 × 2 kernel H.

$$C = \begin{bmatrix} 3 & 8 & 5 & 2 & 20 \\ 15 & 22 & 6 & 9 & 9 \\ 10 & 7 & 3 & 3 & 2 \\ 7 & 4 & 6 & 12 & 19 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 & 5 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 9 \\ 0 & 7 & 0 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & 2 & 0 & 6 & 0 \end{bmatrix}$$

a) Find the four values of the kernel H. [2]

b) What is the padding used? [1]

c) What is the stride used? [1]

d) Why is it so important to keep kernel size small? [1]