Reg. No.: 22BAI4266

Name :

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

Continuous Assessment Test I – August 2024

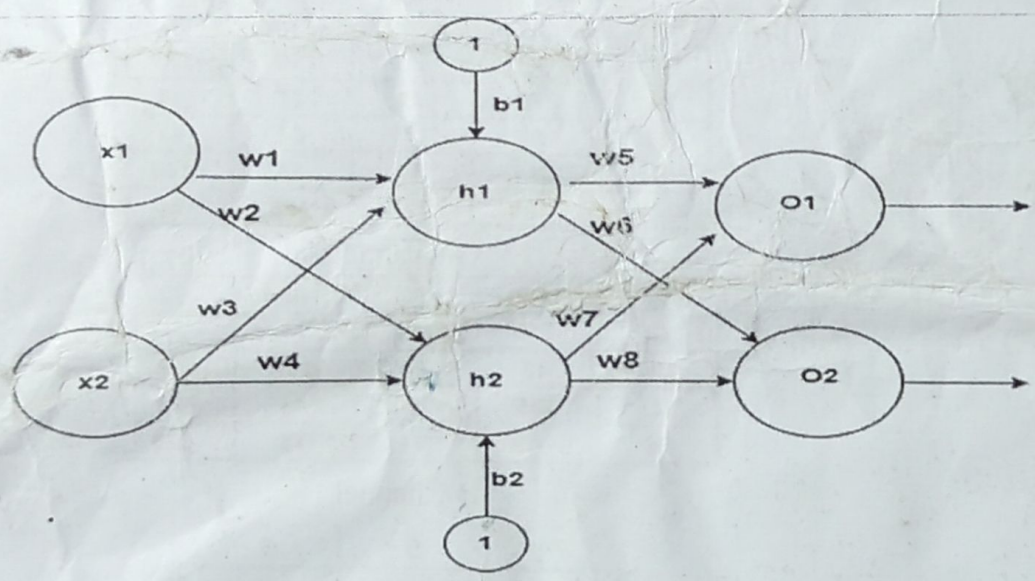
Programme	:	**B.Tech (CSE, AI & ML, MECH, ECE, ECM)**	Semester	:	**Fall 24 - 25**
Course Title	:	**Deep Learning**	Code	:	**BCSE332L**
			Class Nbr(s)	:	**CH2024250100992 CH2024250100984 CH2024250100978**
Faculty (s)	:	**Khadar Nawas K, Pandiaraju V, Rajalakshmi R**	Slot	:	**F1 + TF1**
Time	:	**1½ Hours**	Max. Marks	:	**50**

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Only non-programmable calculator without storage is permitted

Answer all the Questions

1.



15

Consider the above neural network that has 2 neurons each, in the input, hidden and output layers. The weights W, are given below (w1, w2, w3,w4,w5,w6,w7,w8, b1, b2). Use sigmoid activation in the hidden layers and softmax at the output layer. Assume the learning rate to be 1 and error function as MSE.

W = [0.374, 0.920, 0.157, 0.682, 0.498, 0.231, 0.794, 0.546, 0.068, 0.935]

Inputs : x1 = 0.78, x2 = 0.35

Target outputs: O1 = 0.23, O2 = 0.89

Apply Backpropagation algorithm and illustrate each step in detail for both forward and backward passes.

(i) Calculate the outputs at each neuron in the forward path [5]

(ii) Find the error [2]

(iii) Calculate the updated weights for w4 and w8 after back propagation, for one epoch [8]

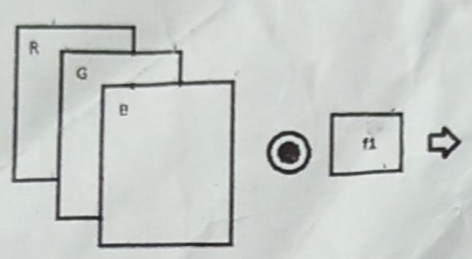
2. a i) Consider a feature map having the size (width x height x depth) 28 x 28 x 250. If you are applying 32 filters of size 1x1 over this feature map, with no padding and a stride of 1, what would be the dimension of output ? Also calculate the number of trainable parameters after applying 1x1 convolution. [4 marks]

ii) Consider the output of Question (i) and find out the size of the feature map after applying Max pooling with a 2*2 filter with stride 1. Write the number of trainable parameters after this pooling operation alone. [3 Marks] 15

b Assume that, a color image is given as input to a CNN and the pixel values for each channel is given below.

i) As shown in Figure, apply a filter f1 on the R, G and B channels of the image and show the obtained feature map. Assume a stride of 2 and no padding. [5 marks]

ii) Apply an average pooling on the obtained output with a filter size of 2x2 and present your output feature maps for all 3 channels separately. [3 marks]



34	78	123	45	67
89	56	101	74	98
12	154	67	199	23
88	34	145	99	101
200	176	88	122	55

R-Channel

77	23	110	56	189
101	76	82	12	88
132	99	178	64	54
95	166	37	88	100
43	198	177	133	74

G-Channel

61	102	90	83	72
18	122	132	97	109
154	48	99	131	66
121	135	190	45	82
76	145	54	78	123

B-Channel

7	5	3
-4	-5	-6
0	8	0

f1

3. You are solving a face recognition-based safety locker design for a bank that uses face image data to help customers to open the locker. You have designed a deep-net classification model using several face images for a specific customer "A" until the model achieves greater than 95% classification accuracy. However, upon deployment of the trained model, it fails to open the safety locker for customer "A" about half the time (50 % failure). 10

a)List the factors you suspect may have led to the mismatch in classification results between the face images used during the designing and the deployment stages. [2 marks]

b) How would you address these issues? Present atleast 3 techniques with a detailed illustration to resolve this problem [6 Marks]

c) Elaborate the procedure for calculating errors in *design and errors in deployment for any* given set of data? [2 marks]

4. You are tasked with developing a neural network to forecast energy demand in a smart grid system. The input features include weather data (temperature, humidity, wind speed), historical energy consumption, and real-time sensor data from various locations across the grid. The network has an input layer, three hidden layers, and an output layer that predicts the energy demand for the next hour.

During the initial training phase, you observe the following issues:

Slow Convergence: The model takes a long time to converge, and even when it does, the performance on the validation set is inconsistent. 10

Gradient Instability: The gradients computed during backpropagation are either too large or too small, leading to unstable updates of the weights.

Compare and contrast the use of Stochastic Gradient Descent (SGD) with momentum, RMSprop, and Adam in addressing the issues observed. Discuss how each of these optimization algorithms could impact the convergence speed, gradient stability, and generalization of your neural network. Which optimizer would you choose to start with, and why?

⇔⇔⇔