U/D/TZ

Reg. No: 0439

Final Assessment Test – April 2026

VIT Vellore Institute of Technology

Course: BCSE332L - Deep Learning

Class NBR(s): 2243 Slot: A1+TA1

Time: Three Hours Max. Marks: 100

- KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE
- DON'T WRITE ANYTHING ON THE QUESTION PAPER

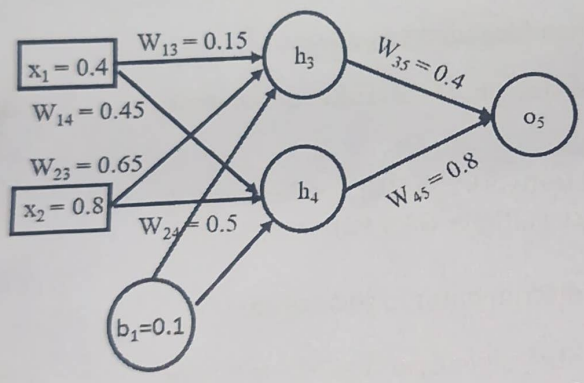
COs	CO Statements
CO1	Understand the methods and terminologies involved in deep neural network, differentiate the learning methods used in Deep-nets.
CO2	Identify and apply suitable deep learning approaches for given application.
CO3	Design and develop custom Deep-nets for human intuitive applications.
CO4	Design of test procedures to assess the efficiency of the developed model.
CO5	To understand the need for Reinforcement learning in real – time problems.

BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Create)

Answer ALL Questions

(10 X 10 = 100 Marks)

1. Perform the forward and backward pass for the network below: **CO1 BL3**



Use a sigmoid activation function and Mean Squared Error (MSE) as a loss function. The targeted output is 1.

2. Explain Mini-batch Gradient Descent and how it differs from batch and stochastic gradient descent. **CO2 BL1**

3. The pixel values of the input image are I. Apply padding of size 1 and perform a convolution process with a stride of 2 using the kernel K. Reduce the dimension of the resultant extracted feature using Max pooling of size 2x2. The values of I and K are: **CO3 BL3**

I=

2	0	2	3	4
3	1	4	2	1
0	2	4	0	2
2	4	3	4	1
1	3	0	1	3

K=

2	2	0
2	1	1
0	1	2

4. For the RNN, perform the forward pass and backward propagation through time with the data given below: **CO3 BL3**

Input sequence length (T): 3 time steps
Input values: $x_1 = 0.5$, $x_2 = 0.8$, $x_3 = -0.9$
Initial hidden state: $h_0 = 0.1$
Input weight: $W_x = 0.5$
Hidden weight: $W_h = 0.6$
Output weight: $W_y = 1.0$
Bias (hidden): $b_h = 0.25$
Bias (output): $b_y = 0.0$
Activation function: tanh
True output labels: $y_1 = 0.6$, $y_2 = 0.7$, $y_3 = -0.6$
Loss function: Mean Squared Error (MSE)
Learning rate (η): 0.1

Show the values of the hidden states, output, loss per time step, and gradients.

5. A system is designed for language translation involving long sentences with complex dependencies. **CO5 BL4**

   i. Explain why long-term dependencies are critical in this task.
   ii. Compare the suitability of standard RNN, LSTM, GRU, and Echo State Networks.
   iii. Discuss optimization strategies for handling long sequences.

6. A startup is developing a facial recognition system, but lacks sufficient diverse training data. **CO3 BL4**

   i. Explain how Generative Adversarial Networks (GANs) can be used to generate synthetic data and suggest suitable GAN variants for this application
   ii. Describe the roles of generator and discriminator in this context.

7. A self-driving car must navigate city traffic, obey signals, and avoid collisions in a highly dynamic environment. **CO4 BL4**

   i. Explain how Deep Reinforcement Learning can be used to train this system.
   ii. Define the state, action, and reward formulation.
   iii. Suggest a suitable Deep Reinforcement Learning algorithm and justify your choice.

8. A financial analyst wants to track the trend of a highly volatile stock price. **CO1 BL2**

   i. Explain how Exponential Weighted Averages can be applied to estimate the moving average.
   ii. Describe how recent values are weighted compared to older values, and analyse how changing β affects responsiveness.
   iii. Discuss the advantages of EWA over simple moving averages.

9.a)	Explain in detail the architecture of BERT, and analyze the advantages of BERT over traditional sequential models like RNNs.	CO1	BL2
	OR		
9.b)	Explain the concept of Gated Recurrent Neural Networks and why they are needed compared to standard RNNs.	CO1	BL2
10.a)	Design a simple CNN architecture for MNIST digit classification (28×28 grayscale images, 10 classes). Your architecture should include at least 2 convolutional layers, pooling layers, and fully connected layers. Specify the dimensions at each layer. Calculate the total number of trainable parameters in your proposed architecture. Explain the role of activation functions (like ReLU) in CNNs and where they should be placed in the architecture.	CO5	BL3
	OR		
10.b)	Distinguish between model-based and model-free reinforcement learning approaches with examples. Discuss the trade-offs between model-based and model-free methods in terms of sample efficiency, computational complexity, and performance.	CO5	BL3

⇔⇔⇔ **U/D/TZ** ⇔⇔⇔