U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

January / February 2025 Semester End Main Examinations

Programme: B.E. **Semester: V**
Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**
Course Code: 24AM5PCDEL **Max Marks: 100**
Course: Deep Learning

B.M.S.C.E. - ODD SEM 2024-25

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

		UNIT - I	***CO***	***PO***	**Marks**
1	a)	Elaborate on the leaning rules of a single layer perceptron.	*CO2*	*PO1*	**05**
	b)	A single-layer perceptron is trained to classify two classes using the Leaky ReLU activation function with $\alpha = 0.01$. The perceptron receives an input vector $x=[2,-3]^\top$ and the weight vector is $w=[0.5,-1.0]^\top$, with a bias $b=0.2$. Find: i. The net input z to the activation function. ii. The output of the Leaky ReLU activation function for the given input.	*CO1*	*PO2*	**10**
	c)	Differentiate between Online and Batch learning.	*CO3*	*PO2*	**05**
		OR			
2	a)	Differentiate between a Single-layer Perceptron and a Multi-layer Perceptron.	*CO2*	*PO2*	**05**
	b)	A Multi-layer Perceptron is designed with the following architecture to solve the XOR problem. • Input layer with 2 neurons (x_1, x_2) • Hidden layer with 2 neurons with sigmoid activation function • Output layer with one neuron operating on sigmoid activation function Given the parameters: Input to Hidden layer weights $\mathbf{W}_{\text{hidden}} = \begin{bmatrix} 10 & 10 \\ -10 & -10 \end{bmatrix}, \quad \mathbf{b}_{\text{hidden}} = \begin{bmatrix} -5 \\ 15 \end{bmatrix}$ Hidden to output layer weights $\mathbf{W}_{\text{output}} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \quad \mathbf{b}_{\text{output}} = -15$ Determine the output of the network for the input combination: $(x_1,x_2)=(0,0)$. Show all intermediate steps.	*CO3*	*PO4*	**10**
	c)	Write the procedural steps of a Perceptron Convergence algorithm.	*CO2*	*PO1*	**05**

			UNIT - II			
	3	a)	Given the true probability distribution (P) and the predicted probability distribution (Q) over three outcomes: <table><tr><th>Outcome (x)</th><th>P(x)</th><th>Q(x)</th></tr><tr><td>1</td><td>0.4</td><td>0.3</td></tr><tr><td>2</td><td>0.4</td><td>0.5</td></tr><tr><td>3</td><td>0.2</td><td>0.2</td></tr></table>Determine the associated cross entropy.	*CO3*	*PO3*	**08**
		b)	List and explain any 2 activation functions of a Deep Neural Network (DNN).	*CO2*	*PO1*	**06**
		c)	Illustrate the functioning of a DNN.	*CO2*	*PO3*	**06**
			OR			
	4	a)	A simple Neural Network is designed with one hidden layer functioning on ReLU (Rectified Linear Unit) activation function and one output layer having Weights of the hidden layer W = [[0.5, -0.5][-0.2, 0.8]] with bias b=[[0.1][0.2]], input vector x=[[-2][3]] and true target label y=1. i. Compute the activations of the hidden layer using the ReLU activation function. ii. Calculate the output from the network if the output layer uses a linear activation. iii. Calculate the Mean Squared Error (MSE) loss.	*CO3*	*PO4*	**08**
		b)	Elaborate on the Vanishing gradient problem and ways to mitigate it.	*CO2*	*PO1*	**06**
		c)	Differentiate between Gradient Descent and Nesterov Accelerated Gradient Descent.	*CO2*	*PO2*	**06**
			UNIT - III			
	5	a)	How does early stopping act as regularizer? Explain.	*CO2*	*PO2*	**10**
		b)	Elaborate the technique of Adaboost optimization.	*CO3*	*PO1*	**10**
			OR			
	6	a)	Are dropouts always helpful during network training? Explain.	*CO2*	*PO2*	**10**
		b)	Elaborate on the technique of RMSprop optimization.	*CO3*	*PO1*	**10**
			UNIT - IV			
	7	a)	Design a Convolutional Neural Network (CNN) with the following specification to classify a gray scale image with the following specifications. Represent the size of the intermediate feature maps precisely on the architectural design. • Input: (64, 64, 1) • Conv1: 16 filters, kernel: 5x5, activation function: ReLU • Pool1: 2x2 Max Pooling • Conv2: 32 filters, kernel: 3x3, activation function: ReLU • Pool2: 2x2 Max Pooling • Flatten • Dense: 64 neurons, activation function: ReLU • Output: 5 neurons, activation function: Softmax	*CO3*	*PO4*	**10**
		b)	What are advantages of padding in a CNN? Explain.	*CO2*	*PO2*	**05**
		c)	Elaborate on backpropagation through convolutions.	*CO1*	*PO1*	**05**
			OR			

B.M.S.C.E. - ODD SEM 2024-25

	8	a)	Consider a Convolutional Neural Network (CNN) with the following configuration: An input image of size 4x4: $\begin{array}{|c|c|c|c|}\hline 1 & 2 & 0 & 3 \\ \hline 4 & 5 & 1 & 0 \\ \hline 2 & 0 & 3 & 1 \\ \hline 1 & 2 & 4 & 0 \\ \hline \end{array}$ A single 2x2 filter (kernel) with weights: $\begin{array}{|c|c|}\hline 1 & -1 \\ \hline 0 & 1 \\ \hline \end{array}$ i. Perform convolution operation on the given input image using the specified filter and stride = 1. Write the resulting feature map. ii. Apply ReLU activation function to the result obtained in (i). Represent the updated feature map in the matrix form. iii. Perform max pooling on the feature map obtained from (ii) using 2*2 pooling window and stride = 1. Determine the resulting feature map.	*CO3*	*PO4*	**10**
		b)	Elaborate on types of padding.	*CO2*	*PO2*	**05**
		c)	Illustrate the concept of convolution as a matrix multiplication.	*CO1*	*PO3*	**05**
			UNIT - V			
	9	a)	Compute the values of input gate and output gate of a LSTM cell by considering the given data: $h_{t-1} = \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix}, x_t = \begin{bmatrix} 0.8 \\ 0.3 \end{bmatrix}$ $W_i = \begin{bmatrix} 0.2 & 0.4 & 0.6 & 0.1 \\ 0.5 & 0.3 & 0.7 & 0.2 \end{bmatrix}, b_i = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$ $W_o = \begin{bmatrix} 0.3 & 0.7 & 0.5 & 0.9 \\ 0.4 & 0.2 & 0.6 & 0.8 \end{bmatrix}, b_o = \begin{bmatrix} 0.05 \\ 0.1 \end{bmatrix}$	*CO3*	*PO4*	**10**
		b)	Illustrate the architecture of Gated Recurrent Units (GRU) cell along with block diagram and related mathematical expressions.	*CO2*	*PO2*	**10**
			OR			
	10	a)	Compute the values of update gate and the Candidate Hidden State $\tilde{h}_t$ of a GRU cell by considering the given data: Current Input (x_t): Previous Hidden State (h_{t-1}): $x_t = \begin{bmatrix} 0.6 \\ 0.2 \end{bmatrix}$ $h_{t-1} = \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix}$ Update Gate Weights and Bias (W_z, b_z): $W_z = \begin{bmatrix} 0.4 & 0.3 & 0.5 & 0.6 \\ 0.7 & 0.2 & 0.4 & 0.3 \end{bmatrix}, \quad b_z = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$ Candidate Hidden State Weights and Bias (W_h, b_h): $W_h = \begin{bmatrix} 0.6 & 0.2 & 0.8 & 0.4 \\ 0.5 & 0.7 & 0.3 & 0.6 \end{bmatrix}, \quad b_h = \begin{bmatrix} 0.05 \\ 0.1 \end{bmatrix}$	*CO3*	*PO4*	**10**

		b)	Illustrate the architecture of Long Short Term Memory (LSTM) cell along with related block diagram and mathematical expressions.	*CO2*	*PO2*	**10**

****

B.M.S.C.E. - ODD SEM 2024-25