U.S.N.										

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

Autonomous Institute Affiliated to VTU

October 2024 Supplementary Examinations

Programme: B.E. **Semester: VI**

Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**

Course Code: 24AM6PCDEL **Max Marks: 100**

Course: Deep Learning

Instructions: 1. Answer any FIVE full questions, choice in UNIT III and UNIT IV only.
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)	i) Illustrate how vanishing gradient can pose problem to the Deep Learning model. ii) Explain how to mitigate it.	*CO2*	*PO2*	**06**
	b)	How properties of ReLU can act as heuristics in avoiding local minima in Deep Neural Networks (DNN)?	*CO1*	*PO1*	**06**
	c)	Differentiate between stochastic gradient descent with momentum and Nesterov accelerated stochastic gradient descent.	*CO1*	*PO1*	**08**
		UNIT - II			
2	a)	A neural network is trained for classification task. It is observed that the training loss continues to decrease, but the validation loss starts to increase after a certain number of epochs. Write i. A meta-algorithm for incorporating early stopping. ii. An algorithm to determine training period.	*CO1*	*PO1*	**08**
	b)	Outline the significance of gradient clipping.	*CO1*	*PO1*	**06**
	c)	Compare RMSProp and RMSProp with Nesterov momentum.	*CO1*	*PO1*	**06**
		UNIT - III			
3	a)	Compute the values of reset gate and the new hidden State h_t of a Gated Recurrent Unit (GRU) cell by considering the given data Weights for the reset gate: $W_r = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.5 & 0.6 & 0.7 & 0.8 \end{bmatrix}$ Biases for the reset gate: $b_r = [0.2, 0.1]$ Input vector: $x_t = [0.5, -0.2]$ Previous hidden state: $h_{t-1} = [0.1, 0.4]$ Candidate hidden state: $\tilde{h}_t \approx \begin{bmatrix} 0.2246 \\ 0.5208 \end{bmatrix}$ Update Gate: $z_t \approx \begin{bmatrix} 0.5646 \\ 0.6630 \end{bmatrix}$	*CO3*	*PO3*	**08**

		b)	Analyse the role of the encoder and decoder in sequence-to-sequence RNN architecture. How are these components structured?	CO2	PO2	**08**
		c)	Illustrate the architecture and working of Multilayer Recurrent Networks.	CO2	PO2	**04**
			OR			
	4	a)	Compute the values of input gate and output gate of a LSTM cell by considering the given data	CO3	PO3	**08**

Weights for the input gate $W_i = \begin{bmatrix} 0.5 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}$ Hidden state $h_{t-1} = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$	Recurrent weights for the input gate $U_i = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}$ Biases for the input gate b_i = [0.2, 0.1]
Input vector $x_t = \begin{bmatrix} 0.5 \\ 0.6 \end{bmatrix}$	Weights for the output gate $W_o = \begin{bmatrix} 0.7 & 0.6 \\ 0.5 & 0.4 \end{bmatrix}$
Recurrent weights for the output gate $U_o = \begin{bmatrix} 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix}$	Biases for the output gate b_o = [0.2, 0.1]

		b)	Explain the steps involved in Modified Backpropagation Through Time (BPTT) algorithm.	CO1	PO1	**06**
		c)	Illustrate Exploding gradient problem of RNN.	CO1	PO1	**06**
			UNIT - IV			
	5	a)	Design a CNN with the following specification to classify images of handwritten digits between 0 and 9. Represent the size of the intermediate feature maps precisely on the architectural design. • **Input**: (32, 32, 1) • **Conv1**: 32 filters, kernel: 3x3, activation function: ReLU • **Pool1**: 2x2 Max Pooling • **Conv2**: 64 filters, kernel: 3x3, activation function: ReLU • **Pool2**: 2x2 Max Pooling • **Flatten** • **Dense**: 128 neurons, activation function: ReLU • **Output**: 10 neurons, activation function: Softmax	CO3	PO3	**10**
		b)	A Convolutional Neural Network is developed for real-time video processing, such as object detection and tracking in live video feeds. Discuss how sparse interactions could be used to balance model complexity and performance in a real-time system.	CO2	PO2	**06**
		c)	Write any four real time applications of adversarial training.	CO1	PO1	**04**
			OR			

	6	a)	Consider a Convolutional Neural Network (CNN) with the following configuration: An input image of size 4x4: $\begin{bmatrix} 1 & 2 & 0 & 3 \\ 4 & 6 & 2 & 1 \\ 1 & 2 & 5 & 3 \\ 0 & 1 & 2 & 4 \end{bmatrix}$ A single 2x2 filter (kernel) with weights: $\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$ i. Perform convolution operation on the given input image using the specified filter. 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. Determine the resulting feature map.	*CO3*	*PO3*	**10**
		b)	Differentiate between convolutional layers, pooling layers, and fully connected layers in Convolutional Neural Network.	*CO1*	*PO1*	**06**
		c)	Differentiate between Tangent distance and Manifold distance.	*CO1*	*PO1*	**04**
			UNIT - V			
	7	a)	Summarize the contrastive divergence algorithm for training Restricted Boltzmann machine.	*CO2*	*PO2*	**06**
		b)	An image denoising project uses a Restricted Boltzmann Machine which has a set of clean and noisy image pairs for training. i. Illustrate the training process of Restricted Boltzmann Machine and explain how it can be used to denoise a new noisy image. ii. what are expected challenges in this application?	*CO1*	*PO1*	**10**
		c)	What are the challenges in applying RBM’s to non-binary data compared to binary data?	*CO2*	*PO2*	**04**

******