MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust®, Mangalore)
Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE, New Delhi
Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

Model Question Paper

Seventh Semester BE Degree Examination

Foundations of Deep Learning

Time: 3 Hours (180 Minutes) **Max. Marks: 100**

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M: Marks, L: RBT (Revised Bloom's Taxonomy) level, C: Course outcomes.

		Module -1	M	L	C
Q1	a.	Compare traditional machine learning approach based on handcrafted features to deep learning approach based on hierarchical representation learning.	08	L2	CO1
	b.	A neuron receives the following pre-activation values: $$[-2, 3, 0, 5, -1]$$ Define the random variable $$X = \text{number of neurons activated by the ReLU function}$$ where $$ReLU(x) = \max(0, x)$$ Compute X.	04	L3	CO2
	c.	Given true label y = [1, 0]and predicted probability yˆ= [0.8, 0.2], compute the cross-entropy loss.	08	L3	CO2
		OR			
Q2	a.	Explain a deep learning architecture that can be used for character recognition	08	L2	CO1
	b.	A sequence of 5-coin flips is used to simulate dropout decisions in a neural network layer. Define the random variable $$X = \text{number of heads}$$ where Head (H) → neuron is masked; Tail (T) → neuron is active. For the experiment outcome: $$TTHHT$$ Determine the value of the random variable X.	04	L3	CO2
	c.	Compute KL divergence between P = [0.6, 0.4], Q = [0.5, 0.5].	08	L3	CO2
		Module- 2			
Q3	a.	Explain the process of vectorizing an image for a machine learning algorithm.	06	L2	CO1
	b.	A neuron has inputs $$x = [1, 2, 3], w = [0.2, 0.5, -0.1].$$ Compute the weighted sum. For a threshold $\theta = 0.5$, apply a step activation function to determine the neuron's output.	06	L3	CO3
	c.	Design a Perceptron Neural Model for XOR operation.	08	L3	CO3

OR					
Q4	a.	Explain how Linear Perceptrons can be expressed as neurons	06	L2	CO1
	b.	Compute the neuron output for the input $$x = [3, 2], w = [0.5, 0.5], b = -2.$$ and determine whether the neuron activates for a threshold $\theta = 0.5$	06	L3	CO3
	c.	A 2–2–1 (2 input neurons, 2 neurons in one hidden layer, 1 output neuron) feed-forward network has: • Input: $x = [1, 2]$ • Hidden layer weights: $$W_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$ • Output layer weights: $$W_2 = [1 \quad 1]$$ Apply forward propagation to compute the output (use linear activation).	08	L3	CO3
Module - 3					
Q5	a.	Explain how stochastic gradient descent enables saddle point avoidance.	04	L2	CO1
	b.	A linear neuron has weight $w = 0.5$. The gradient of the loss with respect to the weight is $\frac{\partial E}{\partial w} = 4$. Apply gradient descent to compute the updated weight for learning rate $\eta = 0.01$	08	L3	CO4
	c.	Given: • hidden neuron activation $a = 0.6$ • error signal $\delta = 0.4$ • learning rate $\eta = 0.1$ Apply backpropagation to compute the updated weight.	08	L3	CO4
OR					
Q6	a.	Describe how minibatch gradient descent overcomes the major pitfalls of stochastic gradient descent.	04	L2	CO1
	b.	Given: • input $x = 1$ • weight $w = 2$ • target $t = 1$ • learning rate $\eta = 0.1$ Apply gradient descent to compute the updated weight of a sigmoidal neuron.	08	L3	CO4
	c.	A two-layer network has output error $\delta^{(2)} = 0.2$and weight $w = 0.5$ connecting hidden and output layer. Apply backpropagation to compute the hidden layer error $\delta^{(1)}$.	08	L3	CO4
Module - 4					

Q7	a.	Compute the new velocity and the updated parameter value for a neural network training using momentum-based optimization, given: Momentum = 0.8 Learning rate = 0.1 Previous velocity = 0.5 Current gradient = 2 Previous parameter value = 5	05	L3	CO5
	b.	A neural network parameter w has initial value $w_0 = 1.0$. The gradient values over three iterations are: $g_1 = 2,\ g_2 = 1,\ g_3 = 2$ Given: 1. Learning rate $\eta = 0.1$ 2. Initial accumulated squared gradient $G_0 = 0$ Apply AdaGrad to compute the updated weight w_3 after three iterations.	07	L3	CO5
	c.	A parameter is trained using Adam optimizer with the following values: ➢ Gradient: $g_t = 2$ ➢ First moment estimate: $m_{t-1} = 1.5$ ➢ Second moment estimate: $v_{t-1} = 3$ ➢ Hyperparameters: $\beta_1 = 0.9, \quad \beta_2 = 0.999, \quad \eta = 0.01$ Apply Adam to compute: 1. Updated m_t 2. Updated v_t 3. Effective step size used to update the parameter	08	L3	CO5
OR					
Q8	a.	Compute the new velocity v_i $and\ the$ updated parameter value θ_i for Nesterov Momentum optimization, where the gradient is evaluated at the look-ahead position instead of the current parameter. the following values are given: ➢ Momentum coefficient $m = 0.8$ ➢ Learning rate $\epsilon = 0.1$ ➢ Previous velocity $v_{i-1} = 0.5$ ➢ Previous parameter value $\theta_{i-1} = 2$ ➢ the gradient evaluated at the look-ahead position = 3.	05	L3	CO5
	b.	Given: ➢ $gradient\ g_t = 4$ ➢ Previous average $E[g^2]_{t-1} = 9$ ➢ Decay rate $\rho = 0.9$ ➢ Learning rate $\eta = 0.01$ Apply RMSProp to compute: 1. Updated the squared gradient moving average $E[g^2]_t$ 2. Effective learning rate used for the weight update	07	L3	CO5

	c.	Compute updated m_t, updated v_t and the effective step size used to update the parameter given that the parameter is trained using Adam optimizer with the following values: ➢ Gradient: $g_t = 3$ ➢ First moment estimate: $m_{t-1} = 1.7$ ➢ Second moment estimate: $v_{t-1} = 4$ ➢ Hyperparameters: $\beta_1 = 0.89, \beta_2 = 0.978, \eta = 0.01$	08	L3	CO5
Module - 5					
Q9	a.	Compute the result of full convolution for the image f(x, y) & kernel h(x, y): **f(x, y)** = [[4, 3], [2, 1]] and **h(x, y)** = [[8, 7], [6, 5]]	10	L3	CO6
	b.	Explain the GoogLeNet Convolutional model with the bock diagram of the inception module. Perform the following steps of the CNN workflow: i) Activation by ReLU ii) Sub sampling by **Max Pooling** 2x2, stride=1 iii) Flattening Layer iv) Dense Layer with Sigmoid function for output calculation. given the convolution result as G(x,y) = [[1, 3, -3, -1], [4, 7, 2, -1], [3, 8, 7, 2]]	10	L3	CO6
OR					
Q10	a.	Compute the result of full cross-correlation for image f(x, y) & kernel h(x, y): **f(x, y)** = [[2, 4], [1, 3]] and **h(x, y)** = [[6, 8], [5, 7]]	10	L3	CO6
	b.	Explain the ResNet CNN architecture with a diagram of the residual connection in ResNet. Perform the following steps of the CNN workflow: i) Activation by ReLU ii) Down sampling by **Average Pooling** 2x2, stride=1 iii) Flatten Layer iv) Dense Layer with Sigmoid function for output calculation. given the cross-correlation result as G(x,y) = [[12, 10, 6, -4], [-5, 7, 3, -2], [9, 4, 11, 5]]	10	L3	CO6

-*****-