102

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2026

B.Tech-IV Semester (ECE/ECS)

COURSE CODE (CREDITS): 25B1WEC432 (3) MAX. MARKS: 35

COURSE NAME: INTRODUCTION TO NEURAL NETWORKS

COURSE INSTRUCTOR: Dr. Nishant Jain MAX. TIME: 2 Hour

Note: (a) All questions are compulsory. (b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems. (c) Calculators are allowed.

Q.No	Question	CO	Marks
Q1	Explain the fundamental trade-off between the depth of frozen layers versus trainable layers when adapting a model to a new domain. How does the size of the target dataset dictate your strategy for choosing which layers to re-train?	CO3	3
Q2	Suppose you are building a feed-forward neural network for a classification task where the input consists of 224x224 RGB images. If the input layer connects to a hidden layer containing 1,024 neurons, calculate the total number of weights in the first layer (assuming a full connection). If you were to reduce the image to grayscale and downsample it to 56x56, how many weights would be saved?	CO3	3
Q3	Explain in detail why standard Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) are structurally unsuited for sequential data. Discuss the three key constraints that these architectures impose, which prevent them from modeling temporal dependencies effectively.	CO4	3
Q4	Given a dataset where input values fluctuate between a minimum of 5 and a maximum of 25, apply the Min-Max scaling formula to normalize an incoming data point of 15. Discuss why normalizing input data is a crucial preprocessing step for gradient-based optimization in deep learning models.	CO4	3
Q5	Consider an RNN unit with an input weight (Wx) of -0.5, a recurrent weight (Wh) of 0.8, and a bias (bx) of 0.2. Calculate the hidden state	CO4	3

	activation (a) for an input (x) of 0.4, assuming the previous hidden state (h_prev) was 0.1. Explain the role of the summation term before applying an activation function.		
Q6	(a) Categorize and differentiate between the four main types of RNN architectures: One-to-One, One-to-Many, Many-to-One, and Many-to-Many. Provide a real-world application example for each. (b) If the hidden state equation is defined as: $H_t = \tanh(W * X_t + W_h * H_{t-1} + b)$, Identify and explain the purpose of each term in this equation.	CO4	2+2=4
Q7	Explain the mechanism of Backpropagation Through Time (BPTT) in the context of sequence modeling. Why is it conceptually different from standard backpropagation in feed-forward neural networks, and what is the primary challenge associated with maintaining hidden state gradients over long sequences?	CO4	4
Q8	Derive the expression for the gradient of a loss function with respect to weights in a multi-layer network. Explain the mechanism behind the vanishing gradient phenomenon and discuss how architectural choices or specific activation functions can mitigate this issue to facilitate effective training in deep models.	CO2	4
Q9	Suppose you are tasked with selecting an architecture for a real-time speech-to-text application where latency is a critical constraint. Compare and contrast the potential performance of a standard RNN, an LSTM, and a GRU in this context.	CO4	4
Q10	Explain the fundamental difference between Supervised and Unsupervised Learning models in the context of high-dimensional data processing. Discuss why Autoencoders are specifically categorized as representation learning models rather than classification or regression tools. How does the concept of 'latent space' act as an information bottleneck to force the model to learn structural patterns?	CO4	4