JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

Make-up Examination-Nov-2025

COURSE CODE (CREDITS):19B1WCI738(3) MAX. MARKS: 25

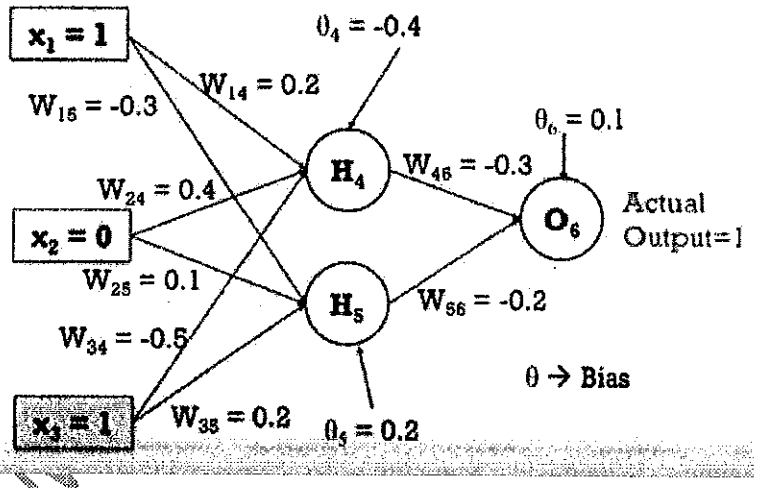
COURSE NAME:INTRODUCTION TO DEEP LEARNING

COURSE INSTRUCTORS: VANI SHARMA MAX. TIME: 1 Hour 30 Minutes

Note:*(a)All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

(c) Use of calculator is allowed.

Q.No	Question	CO	Marks
Q1	Consider the following ANN model with backpropagation algorithm. Weights and biases are given in the figure. The network uses the sigmoid as an activation function. Use the given information to compute the output of each neuron during the forward pass and then calculate the error term for each neuron using the backpropagation algorithm and update the parameters for one iteration using learning rate (η) = 0.1. 	[3]	[7]
Q2	(a) What is Batch Normalization?How does batch normalization help mitigate the vanishing gradient problem? (b) Name any five hyperparameters in terms of an artificial neural network. Give suitable examples.	[3]	[2+2]
Q3	(a) Compare Sigmoid, Tanh, and ReLU activation functions in terms of their mathematical equations, output range, and gradient behavior. (b) What problems arise when using Tanh instead of ReLU in deep neural networks?	[3]	[3+2]

Q4	(a) Suppose you have 5 convolutional kernel of size 7 × 7 with zero padding and stride 1 in the first layer of a convolutional neuralnetwork. You pass an input of dimension 224 x 224 x 3 through this layer. What are the dimensions of the data which the next layer will receive? (b) A convolutional layer has an input size of 64×64×3 and uses 16 filters of size 3×3×3, with stride = 1 and padding = 1.Compute the total number of learnable parameters.	[3]	[2+2]
Q5	Given $y = \sigma(wx + b)$, derive $\frac{dy}{dw}$ in terms of x, y and w. (Hint: use $\sigma'(x) = \sigma(x)(1 - \sigma(x))$.)	[3]	[5]