Name:.....
Roll No:.....

MAR ATHANASIUS COLLEGE OF ENGINEERING (AUTONOMOUS),

KOTHAMANGALAM - 686666

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

B.TECH FIRST SERIES EXAMINATION, AUGUST 2026 (S7 AIML)

AIT 401 FOUNDATIONS OF DEEP LEARNING

Max. Marks: 40 Duration: 1.5 Hours

PART A

Answer all questions. Each question carries 3 marks.

1. Explain the non linear activation functions used in feed forward neural network. Provide examples of scenarios where each might be preferred. (CO1)
2. Given a perceptron with two input features (x1 and x2), weights (w1 and w2), and bias (b), calculate the output of the perceptron for a specific input (x1 = 0.5, x2 = 0.8, w1 = 0.4, w2 = 0.6, b = -0.2) using tanh activation function. Explain the significance of choosing different loss functions.

   (CO1)
3. Consider a two-input neuron that multiplies its two inputs x1 and x2 to obtain the output o. Let L be the loss function that is computed at o. Compute the values of $\partial L/\partial x1$ and $\partial L/\partial x2$ if $\partial L/\partial o = 5$, x1 = 2 and x2 = 3.

   (CO2)
4. Differentiate between Gradient Descent, Stochastic Gradient Descent and Mini-batch Gradient Descent in terms of their working. (CO2)

PART B

Answer all questions. Each question carries 14 marks.

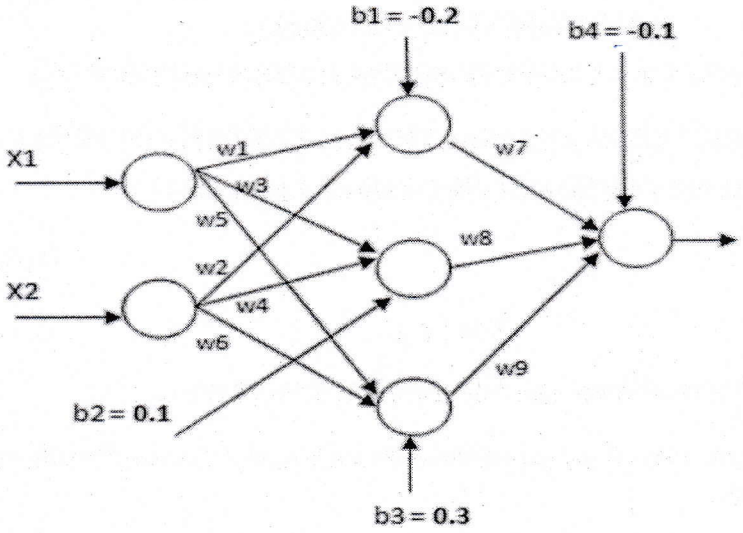
5. Perform one epoch of the back propagation algorithm for the given sample to update the weights and biases using sigmoid activation function and mean squared error loss. Given input vector X=(1,2) , y = (1) , w1 = 0.3, w2 = 0.5, w3 = 0.2, w4 = -0.4, w5= -0.1, w6 = 0.2, w7=0.4,w8 = -0.3, w9 = -0.2, learning rate = 0.1. (CO1)

VISION

Mould competent professionals to serve as valuable resources for society in pursuit of excellence.

MISSION

- Provide quality education tuned to the challenging phases of technology.
- Mould disciplined and Socially Committed engineers capable of professional leadership.
- Promote theoretical and practical understanding, along with critical thinking and logical reasoning.



6. Compare and contrast the variants of GD optimization methods, Momentum, Nestrov , AdaGrad, RMSProp, and Adam. Explain the principles behind these optimization algorithms, including how they adapt learning rates during training. Discuss the scenarios where each of these methods is most effective, and highlight any potential issues or drawbacks. (CO2)