



**B.E DEGREE EXAMINATIONS: APRIL / MAY 2023**

(Regulation 2018)

Seventh Semester

**INFORMATION SCIENCE AND ENGINEERING**

U18ISE0008: Deep Learning

**COURSE OUTCOMES**

- CO1:** Understand the major technology trends driving Deep Learning  
**CO2:** Build, train and apply fully connected deep neural networks  
**CO3:** Effectively use the common neural network techniques  
**CO4:** Appropriately pick hyper parameters

**Time: Three Hours**

**Maximum Marks: 100**

**Answer all the Questions:-  
 PART A (10 x 2 = 20 Marks)  
 (Answer not more than 40 words)**

- |   |     |                   |
|---|-----|-------------------|
| 1. What is binary classification?   | CO1 | [K <sub>2</sub> ] |
| 2. “Convergence depends on the learning rate “ - Justify  | CO1 | [K <sub>4</sub> ] |
| 3. Outline the purpose of non-linear activation function?   | CO2 | [K <sub>2</sub> ] |
| 4. Derive the equation for sigmoid activation function.   | CO2 | [K <sub>2</sub> ] |
| 5. Compare circuit theory and deep learning.  | CO2 | [K <sub>2</sub> ] |
| 6. Write the condition for convergence of a perceptron learning algorithm.  | CO3 | [K <sub>2</sub> ] |
| 7. Assume you have four inputs to a Feed Forward neural network, the first hidden layer also has four neurons, and there are three output classes, what is the dimension of the weight matrix, $W_1$ between the input layer and the first hidden layer, given that there is only one hidden layer? | CO3 | [K <sub>3</sub> ] |
| 8. Differentiate vanishing and exploding gradient.  | CO3 | [K <sub>2</sub> ] |
| 9. Define hyperparameters. List out some of the hyper parameters of Deep Neural Network.  | CO4 | [K <sub>2</sub> ] |
| 10. Why hyper parameter tuning is used in any algorithm?  | CO4 | [K <sub>2</sub> ] |

**Answer any FIVE Questions:-  
 PART B (5 x 16 = 80 Marks)  
 (Answer not more than 400 words)**

- |   |   |     |                   |
|---|---|-----|-------------------|
| 11. a) Discuss in detail about logistic regression cost function.   | 8 | CO1 | [K <sub>2</sub> ] |
| b) Compute the derivatives with a computation graph for the given equation.<br>$J(a,b,c) = 3(a+bc)$ where $a=7$ , $b=3$ and $c = 1$ | 8 | CO1 | [K <sub>3</sub> ] |

- |     |    |  |   |     |                   |
|-----|----|--|---|-----|-------------------|
| 12. | a) | Derive the Gradient Descent equation for neural networks and discuss about the sigmoidal activation function   | 8 | CO2 | [K <sub>2</sub> ] |
|     | b) | Discuss about multiple vectors and derive the vector equation for Forward Propagation for a Shallow Neural Network.  | 8 | CO2 | [K <sub>2</sub> ] |
| 13. | a) | Discuss the characteristics of deep neural networks  | 8 | CO2 | [K <sub>2</sub> ] |
|     | b) | Derive the equation for Forward Propagation and Backward Propagation for a deep neural network.  | 8 | CO2 | [K <sub>3</sub> ] |
| 14. | a) | Describe the hyperparameters used in Adams algorithm and write the algorithm for the same.   | 8 | CO4 | [K <sub>2</sub> ] |
|     | b) | Describe the key components needed to implement a deep neural network. The key components needed are Data, Model , Parameters Algorithm and Error function required. | 8 | CO4 | [K <sub>2</sub> ] |
| 15. | a) | How vectors are initialized in neural networks. Discuss about random initialization method.  | 8 | CO3 | [K <sub>2</sub> ] |
|     | b) | Enumerate the preferred values for Bias and Variance to yield a low Mean Squared Error and explain the trade-off between Bias and Variance of a deep network.        | 8 | CO3 | [K <sub>2</sub> ] |
| 16. | a) | Why can a single layer of perceptron not be used to solve linear inseparable problems? Explain it using an example.  | 8 | CO2 | [K <sub>2</sub> ] |
|     | b) | How normalization is done using a SOFTMAX classifier? Derive the equations with an example.  | 8 | CO4 | [K <sub>3</sub> ] |

\*\*\*\*\*