# 1912E1558

Candidate's Seat No:____

**MSc (AIML) Sem.-3**
**Deep Learning Fundamentals**
**December-2025**

**Time : 3.00 Hours]** **[Max.Marks : 100**

**NOTE : (1) Write both the sections in the separate answer books**
**(2) Figures to the right indicate full marks.**
**(3) Make necessary assumptions wherever necessary.**

## SECTION-I

**Q.1 Answer the following (Attempt Any 2)** **[20]**

(a) What is stride in convolution? Explain how stride, padding, and pooling affect output size and feature extraction. 10

(b) Describe the steps of the backpropagation algorithm for a two-layer network and compute gradients for a sample loss function. 10

(c) What are activation functions? Explain in detail any three advanced activation functions and justify when each should be used. 10

**Q.2 Answer the following questions in details** **[20]**

(a) Compare LSTM, GRU, and Auto-Regressive Networks in terms of architecture, memory gating, and application scenarios. 10

(b) Discuss the steps involved in designing and training a CNN for image recognition 5

(c) Write a note on VGGNet 5

**Q.3 Answer the following** **[10]**

(a) Draw labelled diagrams for RNN, GRU and Bidirectional RNN. Explain how they differ in learning temporal dependencies.

## SECTION-II

**Q.4 Answer the following (Attempt Any 2)** **[20]**

(a) Explain the Bias–Variance decomposition for deep neural networks. Analyze how model complexity influences the bias–variance tradeoff, and discuss strategies to achieve optimal generalization. 10

(b) State and explain the Universal Approximation Theorem. Provide mathematical justification and discuss the practical challenges when using the theorem in deep learning model design. 10

(c) Describe the architecture of a vanilla Multi-Layer Perceptron (MLP). Explain forward pass, activation functions, and capacity control including width, depth, and weight regularization. 10

**Q.5 Answer the following questions in details (Any 2)** **[20]**

(a) Discuss regularization strategies used in deep models including early stopping, dropout, and penalty-based regularization. 10

(b) Write a short note on emerging deep learning trends such as transformers, self-supervised learning, efficient architectures, and model quantization. 10

(c) Outline phases of a complete deep learning workflow including dataset creation, model training, validation, hyperparameter tuning, deployment, and monitoring. 10

P.T.O

E1558 - 2

**Q.6 Answer the following as asked** **[10]**

1. Define Adversarial Training.
2. What is parameter norm penalty?
3. ____ networks use a hierarchical (tree-like) computation structure.
4. Dropout is used to reduce ____.
5. Kernel size in convolution determines ____.
6. Natural gradient methods improve optimization. (True/False)
7. Pooling reduces spatial dimensions of feature maps. (True/False)
8. Gradient clipping helps control exploding gradients. (True/False)
9. Recursive networks cannot process sequential inputs. (True/False)
10. Adam is an adaptive learning rate optimizer. (True/False)

---