

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (2025 - 2026)

Subject: ANN & Deep Learning

Time: 3 Hours

Max. Marks: 100

General Instructions:

IMP: Verify that you have received the question paper with the correct course, code, branch etc.

1. This Question paper comprises of **three Sections -A, B, & C**. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.

2. Maximum marks for each question are indicated on right -hand side of each question.

3. Illustrate your answers with neat sketches wherever necessary.

4. Assume suitable data if necessary.

5. Preferably, write the answers in sequential order.

6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.

SECTION-A

20

1. Attempt all parts:-

(1.1) How many layers does a Multilayer Perceptron typically have (CO1, K2)

1

- (a) 1
- (b) 2
- (c) 3
- (d) 4 or more

(1.2) Artificial neural network used for- (CO1, K3)

1

- (a) Pattern Recognition
- (b) Classification
- (c) Clustering
- (d) All of these

(1.3) The main purpose of a sparse autoencoder is (CO2, K2)

1

- (a) To learn a compressed representation of the input data
- (b) To remove noise from the input data
- (c) To increase the dimensionality of the input data
- (d) To produce a sparse output representation

(1.4) The computational complexity of singular value decomposition is (CO2, K2)

1

- (a) $O(n)$
- (b) $O(n^2)$
- (c) $O(n^3)$
- (d) $O(n \log n)$

(1.5) Low Bias refers to (CO3, K2)

1

- (a) accuracy
 - (b) confusion matrix
 - (c) underfitting
 - (d) overfitting
- (1.6) Which of the following is NOT an advantage of Batch Normalization? (CO3, K2) 1
- (a) It can reduce overfitting in the model
 - (b) It can improve the accuracy of the model on the validation set
 - (c) It can speed up the training process by reducing the number of epochs needed
 - (d) It can reduce the complexity of the model
- (1.7) CNN has more layers than ANN (CO4, K2) 1
- (a) FALSE
 - (b) TRUE
 - (c) Partially true
 - (d) None of these
- (1.8) Sigmoidal activation function is generally used at (CO4, K3) 1
- (a) hidden layers
 - (b) output layer
 - (c) both A & B
 - (d) none of these
- (1.9) In Backpropagation Through Time (BPTT), what is the role of the loss function? (CO5, K2) 1
- (a) To compute the gradients for each layer of the recurrent neural network
 - (b) To determine the optimal number of time steps to use for the BPTT algorithm
 - (c) To measure the performance of the model and provide feedback for the weight updates
 - (d) To determine the learning rate for the weight updates in the BPTT algorithm
- (1.10) Which of the following is an advantage of using LSTM networks over traditional RNNs? (CO5, K2) 1
- (a) They are less prone to the vanishing gradient problem.
 - (b) They are faster to train.
 - (c) They require less memory to store parameters
 - (d) They are more resistant to overfitting.

2. Attempt all parts:-

- (2.1) Distinguish between Supervised and Unsupervised Learning. (CO1, K4) 2
- (2.2) How does regularization affect the performance of Gradient Descent? (CO2, K2) 2
- (2.3) Discuss Softmax layer in deep learning. (CO3, K2) 2
- (2.4) List down the hyperparameters of a Pooling Layer.(CO4, K1) 2
- (2.5) Evaluate common applications of GRUs. (CO5, K5) 2

SECTION-B

30

Answer any one of the following:-

(3.1)	Describe the linear and nonlinear activation functions used in Artificial Neural Networks. (CO1, K2)	6
(3.2)	Explain (1) binary sigmoidal activation function, (2) bipolar sigmoidal activation function.(CO1, K2)	6
Answer any one of the following:-		
(3.3)	Examine the key steps involved in performing PCA on a dataset and interpret the results. (CO2, K2)	6
(3.4)	Apply sparse autoencoders for anomaly detection and evaluate their benefits. (CO2, K3)	6
Answer any one of the following:-		
(3.5)	Analyze why ReLU is preferred over sigmoid in CNN. (CO3, K4)	6
(3.6)	Explain how early stopping improves performance. (CO3, K2)	6
Answer any one of the following:-		
(3.7)	What is LeNet ? Explain it with its architecture? (CO4, K2)	6
(3.8)	An input image has been converted into a matrix of size 12 X 12 along with a filter of size 3 X 3 with a Stride of 1. Determine the size of the convoluted matrix. (CO4, K2)	6
Answer any one of the following:-		
(3.9)	Compare GRU with a long short-term memory (LSTM) network. (CO5, K4)	6
(3.10)	Compare BPTT with the standard backpropagation algorithm used in feedforward neural networks. (CO5, K3)	6
<u>SECTION-C</u>		50
4. Answer any <u>one</u> of the following:-		
(4.1)	Draw and explain the Multilayer ANN model through an appropriate example. (CO1, K2)	10
(4.2)	Explain Multilayer ANN systems? (CO1, K2)	10
5. Answer any <u>one</u> of the following:-		
(5.1)	What do you understand by regularization in autoencoder and also discuss denoising autoencoder in brief. (CO2, K2)	10
(5.2)	Describe gradient descent with figure and also write its advantage and disadvantage.(CO2, K2)	10
6. Answer any <u>one</u> of the following:-		
(6.1)	Illustrate the greedy layerwise pre-training in deep learning. (CO3, K2)	10
(6.2)	Write short note on L2 and L1 regularization. (CO3, K2)	10
7. Answer any <u>one</u> of the following:-		
(7.1)	Explain the working of ZFNet with a suitable architecture. (CO4, K2)	10
(7.2)	What are the advantages and disadvantages of DenseNet over AlexNet? (CO4, K2)	10
8. Answer any <u>one</u> of the following:-		
(8.1)	Compare recurrent neural network, long short-term memory network and Gated Recurrent Unit (GRU) with diagram. (CO5, K2)	10
(8.2)	How does the vanishing gradient problem arise in BPTT, and what are some	10

techniques for mitigating it? (CO5, K2)

REG_JAN_JUNE_2026