

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

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) | In artificial Neural Network interconnected processing elements are called (CO1, K1) | 1 |
| | <ul style="list-style-type: none"> (a) nodes or neurons (b) weights (c) axons (d) soma | |
| (1.2) | The need of biological neural networks is (CO1, K1) | 1 |
| | <ul style="list-style-type: none"> (a) to solve tasks like machine vision & natural language processing (b) to apply heuristic search methods to find solutions of problem (c) to make smart human interactive & user friendly system (d) all of the mentioned | |
| (1.3) | What is the goal of Gradient Descent? (CO2, K2) | 1 |
| | <ul style="list-style-type: none"> (a) To maximize the objective function (b) To minimize the objective function (c) To improve the accuracy of the model (d) To reduce the number of parameters in the model | |
| (1.4) | Principal components calculated in PCA.....(CO2,K1) | 1 |
| | <ul style="list-style-type: none"> (a) by multiplying the original data matrix by its transpose (b) by finding the eigenvectors and eigenvalues of the covariance matrix (c) by calculating the mean of each variable in the dataset (d) by using a clustering algorithm | |

- (1.5) How does Early Stopping work? (CO3, K2) 1
- (a) by stopping the training process when the validation error stops improving
 - (b) by increasing the number of iterations during training
 - (c) by decreasing the learning rate during training
 - (d) by adding more layers to the neural network
- (1.6) What is the impact of increasing training set size on the bias-variance tradeoff? (CO3, K2) 1
- (a) It increases bias and decreases variance
 - (b) It increases variance and decreases bias
 - (c) It increases both bias and variance
 - (d) It decreases both bias and variance
- (1.7) Which of the following is a subset of machine learning? (CO4,K1) 1
- (a) Numpy
 - (b) SciPy
 - (c) Deep Learning
 - (d) none of the above
- (1.8) CNN is mostly used when there is an- (CO4, K1) 1
- (a) structured data
 - (b) unstructured data
 - (c) both A & B
 - (d) none of the above
- (1.9) What is the purpose of the hidden state in an RNN? (CO5, K2) 1
- (a) To store information from previous time steps
 - (b) To store information from the current time step
 - (c) To store information from the next time step
 - (d) To store information from multiple time steps
- (1.10) 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

2. Attempt all parts:-

- (2.1) How are Artificial Neural Networks different from Normal Computers? (CO1,K2) 2
- (2.2) What is Gradient Descent and how does it work? (CO2,K3) 2
- (2.3) Define type I error with suitable example? (CO3,K2) 2
- (2.4) List down the hyperparameters of a Pooling Layer. (CO4,K2) 2
- (2.5) How does the input gate work in an LSTM cell? (CO5,K3) 2

SECTION-B

30

Answer any one of the following:-

- (3.1) How many types of the artificial neural network used in machine learning? Explain any two in brief? (CO1, K2) 6
- (3.2) Calculate the net input for $x_1=0.2$, $x_2=0.6$, and bias $b=1$ with weight 0.3. Assume $w_1=w_2=1$. (CO1, K3) 6

Answer any one of the following:-

- (3.3) List down the advantages and disadvantages of gradient descent? (CO2, K1) 6
- (3.4) Write the short note on momentum based gradient descent. (CO2, K4) 6

Answer any one of the following:-

- (3.5) How the weight updation process takes place in CNN? Explain with suitable diagram. (CO3, K3) 6
- (3.6) If true positive = 10, false positive = 20, true negative = 15, and false negative = 10, calculate the value of precision and recall. (CO3, K4) 6

Answer any one of the following:-

- (3.7) Elaborate Softmax and ReLU activation functions in brief. (CO4, K2) 6
- (3.8) How the weights of layers gets changed? Explain in detail. (CO4, K3) 6

Answer any one of the following:-

- (3.9) How does BPTT differ from the standard backpropagation algorithm used in feedforward neural networks? (CO5, K3) 6
- (3.10) What strategies can be employed to tackle the vanishing gradient issue in Recurrent Neural Networks (RNNs)? (CO5, K2) 6

SECTION-C 50

4. Answer any one of the following:-

- (4.1) How we employ deep learning to create an expert system? Please provide a detailed explanation. (CO1, K2) 10
- (4.2) Examine five significant applications of Deep Learning. Explain each in detail. (CO1, K3) 10

5. Answer any one of the following:-

- (5.1) Describe gradient descent with figure and also write its advantage and disadvantage. (CO2, K2) 10
- (5.2) Decompose the given matrix $\begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$ using singular value decomposition. (CO2, K3) 10

6. Answer any one of the following:-

- (6.1) Write short note on the following: (i) ReLU activation function, (ii) sigmoid function, (iii) tanh (iv) linear activation function (v) Leaky ReLU activation function. (CO3, K2) 10
- (6.2) Illustrate and analyze the process of early stopping and batch normalization in deep learning. (CO3, K3) 10

7. Answer any one of the following:-

- (7.1) How does ZFNet work? Explain with suitable architecture. (CO4, K2) 10
- (7.2) Elaborate the working of DenseNet with suitable diagram. (CO4, K2) 10

8. Answer any one of the following:-

- | | | |
|-------|---|----|
| (8.1) | Describe in detail all the gates used in long short-term memory (LSTM) network.
(CO5,K2) | 10 |
| (8.2) | Draw and explain the working of RNN in detail. (CO5,K4) | 10 |

REG_JAN_JUNE_2026