Roll. No:

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

M.Tech(Integrated)

SEM: VII - THEORY EXAMINATION (2024 - 2025)

Subject: Programming for Data Analytics

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-a.   Tick the correct choice is a data structure in Pandas?(CO1,K1)   1

      (a)   Set

      (b)   Tuple

      (c)   Series

      (d)   Dictionary

1-b.   Select the correct choice for date series element in Pandas?(CO1,K2)   1

      (a)   Using .iloc[] accessor function

      (b)   Using indexing operator [] and passing in the name of the column

      (c)   Using .loc[] accessor function

      (d)   Using indexing operator [] and passing in the label or integer location of the element

1-c.   A _____ is an R-object which can contain many different types of elements   1
            inside it.(CO2,K1)

      (a)   List

      (b)   Vector

      (c)   Factor

      (d)   Matrix

1-d.   CRAN stands for (CO2,K1)   1

      (a)   Comprehensive R Architecture Network

      (b)   Comprehensive R Archive Network

(c)   Comprehensive R Automatic Network

(d)   Complete R Archive Network

1-e.   Name the .exit command do?(CO3,K1)   1

(a)   It exits the sql prompt

(b)   It exits the sqlite prompt

(c)   It exits the PL/SQL prompt

(d)   None of the above

1-f.   Name the .header(s) onoff command do?(CO3,K1)   1

(a)   Display of headers are turned on

(b)   Display of headers are turned off

(c)   Display of headers are turned on or off

(d)   None of the above

1-g.   Correct choice for TensorFlow?(CO4,K1)   1

(a)   A programming language

(b)   A machine learning library

(c)   A deep learning framework

(d)   An operating system

1-h.   Name Out Common programming languages are used by TensorFlow?(CO4,K1)   1

(a)   Python only

(b)   Java only

(c)   C++ only

(d)   Python, Java, and C++

1-i.   List of the following is NOT a commonly used activation function in deep learning?(CO5,K2)   1

(a)   Sigmoid

(b)   Tanh

(c)   Linear

(d)   None of these

1-j.   Select the right one of the following is NOT a type of layer commonly used in deep neural networks?(CO5,K2)   1

(a)   Convolutional layer

(b)   Pooling layer

(c)   Linear layer

(d)   None of these

2. Attempt all parts:-

2.a.   Write short notes on Pandas with Common suitable library?(CO1,K2)   2

2.b.   Define vector with suitable command.(CO2,K1)   2

2.c.   Which command in MongoDB is used to see collections?(CO3,K2)   2

| 2.d. | Briefly mention role of Session in TensorFlow .(CO4,K2) | 2 |
| 2.e. | Short note on deep learning Association with Data Analytics?(CO5,K2) | 2 |

SECTION-B   30

3. Answer any <u>five</u> of the following:-

3-a.	Differentiate between Series and Array and briefly explain the concept of Series in Pandas.(CO1,K2)	6
3-b.	Illustrate the commands to create program for Pandas Series from a list of numbers and then multiply each element in the Series by 2.(CO1,K3)	6
3-c.	Explain the use of factor () with example.(CO2,K2)	6
3-d.	Specify concept of mean(is.na(x)) tell you about a vector x? What about sum(!is.finite(x)) ?(CO2,K2)	6
3-e.	Explain the concept of pipeline in the MongoDB aggregation framework.(CO3,K2)	6
3-f.	Explain the process of building a simple linear regression model using TensorFlow.(CO4,K3)	6
3-g.	Discuss the challenges of training deep neural networks on large-scale datasets and some techniques to address them.(CO5,K2)	6

SECTION-C   50

4. Answer any <u>one</u> of the following:-

| 4-a. | Steps for some common measures of central tendency and variability, and how are they used in descriptive statistics?(CO1,K3) | 10 |
| 4-b. | Differentiate between a confidence interval and a p-value? And also expain some assumptions that must be met for the p-test to be valid.(CO1,K2) | 10 |

5. Answer any <u>one</u> of the following:-

| 5-a. | Differentiate between a bar-chart and a histogram? Where would you use a bar-chart and where would you use a histogram?(CO2,K2) | 10 |
| 5-b. | Explain Predictive analysis process in R?(CO2,K2) | 10 |

6. Answer any <u>one</u> of the following:-

| 6-a. | Step down the procedure to create a database in MongoDB and mention various common databases.(CO3,K2) | 10 |
| 6-b. | Differentiate between Relational Data Base with NOSQL and mention popular NoSQL Databases.(CO3,K2) | 10 |

7. Answer any <u>one</u> of the following:-

| 7-a. | Explain concept to visualize the filters learned by a convolutional layer in a CNN using Python?(CO4,K2) | 10 |
| 7-b. | Illustrate that some common activation functions used in CNNs?, Explain in details.(CO4,K2) | 10 |

8. Answer any <u>one</u> of the following:-

| 8-a. | Specify Deep neural networks commonly used for unsupervised learning, and | 10 |

write short notes on Autoencoders and Generative Adversarial networks?(CO5,K2)

| 8-b. | Explain the Autoencoders and Variational Encoders with neat diagram structures.(CO5,K2) | 10 |

.