Printed Page:- **Subject Code**:- **ACSAI0617**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (20 ..... - 20.....)

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. Apply the method that calculates covariance.(CO1,K1) 1

(a) df.cov()
(b) df.corr()
(c) df.var()
(d) df.mean()

1-b. Choose the function to compute the mean of a NumPy array.(CO1,K1) 1

(a) np.mean()
(b) np.median()
(c) np.std()
(d) np.var()

1-c. What will be the output of log (-5.8) when executed on R console?(CO2,K2) 1

(a) NAN
(b) NA
(c) Error
(d) 0.213

1-d. How will you check if an element is present in a vector?(CO2,K1) 1

(a) match()
(b) dismatch()
(c) mismatch()

(d) search()

1-e. Which of the following is/are SQLite command(s)?(CO3,K1) 1

(a) DDL
(b) DML
(c) DQL
(d) All of the above

1-f. Which one of the following is not a comparison query operator in MongoDB?(CO3,K2) 1

(a) $eq
(b) $gte
(c) $nor
(d) None of these

1-g. Component used to observe model training flow(CO4,K1) 1

(a) TensorFlow Playground
(b) TensorFlow Quickstart
(c) TensorFlow Estimators
(d) TensorBoard Visualization

1-h. Role of Word Vectors in natural language understanding(CO4,K2) 1

(a) Encodes words numerically
(b) Deletes words
(c) Changes font of words
(d) Aligns words

1-i. Common technique to handle class imbalance in classification tasks(CO5,K1) 1

(a) Oversampling
(b) Undersampling
(c) SMOTE
(d) All of the above

1-j. Technique used to automatically adjust learning rate during training(CO5,K2) 1

(a) Learning Rate Decay
(b) Gradient Descent
(c) Momentum
(d) Adagrad

2. Attempt all parts:-

2.a. Describe the steps for performing a t-test in Python using Pandas.(CO1,K2) 2

2.b. Describe the use of str_split() in string processing with a basic illustration.(CO2,K2) 2

2.c. State the function of fetchall() in data retrieval from a database cursor.(CO3,K2) 2

2.d. Describe how TensorFlow simplifies the process of building and training neural 2

.

networks.(CO4,K2)

2.e.	Primary use of Keras in the context of advanced deep learning(CO5,K2)	2

SECTION-B 30

3. Answer any five of the following:-

3-a.	Demonstrate the steps to create a scatter plot using Matplotlib in Python.(CO1,K2)	6
3-b.	Describe the process of using regular expressions to identify patterns in text data.(CO1,K2)	6
3-c.	Name the function used for horizontal merging of data frames in R and show its usage through an example.(CO2,K2)	6
3-d.	Demonstrate the plotting behavior of geom_col() and distinguish it from geom_bar() using one example for each.(CO2,K3)	6
3.e.	What are the aggregate functions of MongoDB?(CO3,K2)	6
3.f.	Explore applications of RNNs in language modeling, speech recognition, and time series prediction.(CO4,K2)	6
3.g.	Discuss the concept of autoencoders and elaborate on their role in learning efficient data representations.(CO5,K2)	6

SECTION-C 50

4. Answer any one of the following:-

4-a.	Perform data filtering and selection operations on a Pandas DataFrame with sample data.(CO1,K3)	10
4-b.	Apply NumPy functions for matrix multiplication and statistical computations.(CO1,K3)	10

5. Answer any one of the following:-

5-a.	Explain the finction of filter(), arrange() and mutate() with suitable examples(CO2,K2)	10
5-b.	Apply arithmetic operations on two numeric vectors v1 and v2. Multiply vectors, cube elements of v1, convert v2 to integers using two methods, and create sequences using seq() and colon operator.(CO2,K3)	10

6. Answer any one of the following:-

6-a.	What do you understand by NoSQL databases? Is MongoDB a NoSQL database? Explain.(CO3,K2)	10
6-b.	Differentiate Between $in and $nin ,$lt and $lte, $not and $nor, $eq and $ne.(CO3,K3)	10

7. Answer any one of the following:-

7-a.	Demonstrate the implementation of model parallelism and data parallelism using Distributed TensorFlow.(CO4,K4)	10
7-b.	Define a CNN architecture using TensorFlow/Keras for image classification. Include convolutional layers, pooling layers, activation functions, and fully connected layers. (CO4,K2)	10

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

8-a.	Explore the concept of Disentangled Representation GANs, detailing their architectures, loss functions, and usage in separating factors of variation.(CO5,K2)	10
8-b.	Analyze the challenges and solutions related to exploration-exploitation balance in Deep Reinforcement Learning algorithms.(CO5,K4)	10

.