

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

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (2022-2023)

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. | How do you create a Series in Pandas?(CO1) | 1 |
| | (a) Using pd.Series() function and passing in a Python list or NumPy array | |
| | (b) Using pd.DataFrame() function and passing in a Python dictionary or a list of lists | |
| | (c) Using pd.Series() function and passing in a Python dictionary | |
| | (d) Using pd.DataFrame() function and passing in a Python list or NumPy array | |
| 1-b. | Which method can be used to drop rows with missing values in pandas?(CO1) | 1 |
| | (a) df.fillna() | |
| | (b) df.dropna() | |
| | (c) df.replace() | |
| | (d) df.interpolate() | |
| 1-c. | CRAN stands for(CO2) | 1 |
| | (a) Comprehensive R Architecture Network | |
| | (b) Comprehensive R Archive Network | |

(c) Comprehensive R Automatic Network

(d) Complete R Archive Network

- 1-d. _____ function is used for reading the .csv file in R language.(CO2) 1
- (a) write.csv()
 - (b) read.csv()
 - (c) let.csv()
 - (d) table.csv()
- 1-e. The \$rename operator logically performs a _____ of both the old name and the new name.(CO3) 1
- (a) \$unset
 - (b) \$set
 - (c) \$Nested
 - (d) None of the above
- 1-f. NoSQL stands for(CO3) 1
- (a) Not Only SQL
 - (b) Not a SQL
 - (c) New SQL
 - (d) None of the above
- 1-g. Which of the following is a high-level API for building deep learning models in TensorFlow?(CO4) 1
- (a) TensorFlow.js
 - (b) TensorFlow Hub
 - (c) TensorFlow Lite
 - (d) TensorFlow Keras
- 1-h. Which of the following layers is commonly used in a CNN for feature extraction?(CO4) 1
- (a) Dense layer
 - (b) Pooling layer
 - (c) Dropout layer
 - (d) Activation layer
- 1-i. Which of the following is a technique used to prevent overfitting in deep learning models?(CO5) 1
- (a) Dropout

- (b) Batch normalization
 - (c) Data augmentation
 - (d) Normalization
- 1-j. Which of the following is a commonly used pre-processing technique for input data in GANs?(CO5) 1
- (a) Data normalization
 - (b) Data augmentation
 - (c) Transfer learning
 - (d) Transfer loss

2. Attempt all parts:-

- 2.a. What is sensitivity in a confusion matrix?(CO1) 2
- 2.b. How to Use na.rm in R?(CO2) 2
- 2.c. What are Databases in MongoDB?(CO3) 2
- 2.d. What is a convolutional layer in a CNN?(CO4) 2
- 2.e. What is an activation function?(CO5) 2

SECTION B

30

3. Answer any five of the following:-

- 3-a. What is a Pandas Series and how is it different from a NumPy array?(CO1) 6
- 3-b. What is sensitivity analysis, and how is it used in decision making?(CO1) 6
- 3-c. Compare and contrast geom_jitter() with geom_count().(CO2) 6
- 3-d. What are the different data structures in R? Briefly explain about them.(CO2) 6
- 3.e. Explain Sharding and Aggregation in MongoDB?(CO3) 6
- 3.f. How does transfer learning work in TensorFlow, and what are some common use cases for it?(CO4) 6
- 3.g. Discuss the challenges of training deep neural networks on large-scale datasets and some techniques to address them.(CO5) 6

SECTION C

50

4. Answer any one of the following:-

- 4-a. Describe the term: (a) Mean (b) Median (C) Standard Deviation (d) Mode (e) Variance(CO1) 10
- 4-b. Explain (a) What is data wrangling, and why is it important in data analysis? (b) How can you clean data in Pandas, and what are some common data cleaning tasks?(CO1) 10

5. Answer any one of the following:-

- 5-a. Create following vectors v1(1, 2, 13, 5, 10) and v2(1, 56, 78, 23, 7). and do as directed (CO2) 10
a) Multiply
b) Take cube of each element of v1
c) Create a vector using seq() method
- 5-b. Explain Predictive analysis process in R?(CO2) 10

6. Answer any one of the following:-

- 6-a. What are the several reasons why you might choose to use SQLite in your project?(CO3) 10
- 6-b. Explain types of NoSQL databases. How is NoSQL different from RDBMS?(CO3) 10

7. Answer any one of the following:-

- 7-a. What are some common activation functions used in CNNs?, Explain in details.(CO4) 10
- 7-b. Explain the working of RNN model and compare with CNN?(CO4) 10

8. Answer any one of the following:-

- 8-a. How can deep neural networks be used for reinforcement learning, and what are some common applications in this field?(CO5) 10
- 8-b. Explain the working of the GAN model and also write applications.(CO5) 10