CCMTE/D-23   24058
DATA SCIENCE USING PYTHON
MT-CSE-20-13(ii)

Time : Three Hours]   [Maximum Marks : 75

Note : Attempt *Five* questions in all, selecting *one* question each from Unit I to Unit IV. Q. No. **1** is compulsory. All questions carry equal marks.

Compulsory Question

1. (a) How are big data and data science related ?

   (b) Differentiate between continuous and categorical datasets.

   (c) What is variance and deviation ? Give example.

   (d) What is the CLT for sample proportions ?

   (e) Define Planar and Retinal Variables.

   (f) Describe data encoding process.

   (g) Differentiate between NumPy and SciPy.

   (h) What is monkey patching in Python ?   **15**

Unit I

2. (a) Outline and explain data science process. Highlight the role of data scientist in this process.

(b) What are data science toolkit ? How do these toolkit support data science applications ? **8+7=15**

3. (a) What are the different sources of data ? Write API to retrieve data from social platform.

(b) Write the different strategy to fix the faulty data. How is quality data stored and managed ? **8+7=15**

Unit II

4. (a) What is the purpose of central tendency ? How do you find the central tendency ? Discuss the for measures of central tendency.

(b) What is distributive property of integers ? Distributive property of multiplication over addition and over subtraction. **8+7=15**

5. Differentiate between supervised, unsupervised and reinforcement machine learning algorithms. Discuss :
(i) Different kernel functions in SVM
(ii) Gaussian and Bernoulli Naïve Bayes. **8+7=15**

Unit III

6. What is data visualization ? Explain different types of data visualization. Explore the recent trends of data collection and analysis techniques. **15**

7. (a) What is visual encoding ? Why is it important ? How do you choose appropriate visual encoding ?

(b) Write a note on methods used for application development in data science. **8+7=15**

Unit IV

8. (a) Write the scope and importance of Python. Differentiate between :
(i) list and tuples,
(ii) modules and packages.

(b) Illustrate a Python script to explore break, continue and pass statement. **8+7=15**

9. (a) How to read and write from a file in Python ? Discuss user-defined exceptions in Python. State example.

(b) Write a Python program to eliminate repeated lines from a file. **8+7=15**