

THIRD SEMESTER B.Sc. DEGREE EXAMINATION, NOVEMBER 2024
(Regular/Improvement/Supplementary)
HONOURS IN MATHEMATICS
GMAH3B14T: ADVANCED PYTHON PROGRAMMING

Time: 3 hours

Maximum Marks: 80

Part A. Answer *all* the questions. Each carries *one* mark.

Choose the correct answer.

1. In NumPy, the shape attribute provides
A) Array dimensions B) Array index C) Type D) Location
2. The attribute used to find the data type of a NumPy array is
A) type B) dtype C) dtype D) datatype
3. Which among the following is not a supervised machine learning algorithm?
A) Naive Bayes B) SVM C) K-means D) Decision tree
4. ____ prints the current row index of the DataFrame
A) DataFrame[0] B) DataFrame[1] C) rowIndex D) index
5. The ____ method returns the count of each unique values of a series.
A) value_counts() B) distinct() C) count() D) get_counts()

Fill in the Blanks.

6. ____ is a type conversion method in Pandas.
7. The ____ method is used to load a CSV file to pandas DataFrame.
8. The default data type of NumPy array is _____.
9. ____ is an application of machine learning.
10. The fit() method in scikit Learn is used for _____.

(10 × 1 = 10 Marks)

Part B. Answer any *eight* questions. Each question carries *two* marks.

11. Discuss how groupby() method works in Pandas.
12. What is the difference between indexing and slicing in NumPy?
13. What is linear regression model?
14. Discuss how to encode categorical data in scikit Learn.
15. Define data analytics.

(PTO)

16. What is supervised learning?
17. Differentiate `iloc()` and `loc()` of Pandas DataFrame.
18. What is over fitting? How can you avoid it?
19. Define reinforcement learning.
20. How do you create a random 2-D array using NumPy?

(8 × 2 =16 Marks)

Part C. Answer any *six* questions. Each carries *four* marks.

21. Define clustering. What are the different types of clustering?
22. What are the different ways to create a DataFrame in Pandas? Explain.
23. Write a note on model selection.
24. Write a NumPy program to count number of zeros and non zero values in an array.
25. What is a series? Write the code to retrieve unique values from a column in Pandas DataFrame.
26. Explain the operations that can be performed on NumPy arrays.
27. List and explain the various steps involved in data analytic process.
28. Write a note on data aggregation and summarization with respect to Pandas.

(6 × 4 = 24 Marks)

Part D. Answer any *two* questions. Each carries *fifteen* marks.

29. How can you create a plot in NumPy? Write the code to plot a line graph.
30. What is data visualization? Explain with an example how bar plots, and histograms are plotted using Seaborn.
31. What is unsupervised learning? Explain various learning techniques involved in unsupervised learning?

(2 × 15 = 30 Marks)