Reg. No ..... Name **24UP3003**

END SEMESTER EXAMINATION - OCTOBER 2024

SEMESTER 7: INTEGRATED M.Sc. PROGRAMME COMPUTER SCIENCE - DATA SCIENCE

COURSE: **21UP7CRMCP23 : ADVANCED PYTHON PROGRAMMING FOR DATA SCIENCE**

(For Regular - 2021 Admission)

Time: Three Hours Max. Weightage: 30

PART A
Answer any 8 Questions

1. Define Dataset. List any four sources of dataset.
2. State the purpose of dropout layer in a neural network.
3. Write a simple code snippet that would generate random numerical data having four features.
4. Write any five statements that describes various ways of reading data from Excel files.
5. Differentiate between continuous and categorical data in the context of a dataset.
6. List any five python modules that would be helpful in scientific computing.
7. Describe TensorFlow sessions.
8. Write the statement that imports `matplotlib` module. Also, list any four types of plots supported by `matplotlib` module along with the methods used to create the plot.
9. Define Reinforcement learning.
10. Define lambda function, along with an example.

(1 x 8 = 8 Weight)

PART B
Answer any 6 Questions

11. With a simple figure, explain the components of Matplotlib. Also, write a simple program that plots a line in the 2D space.
12. With sample programs, explain how Python can be used to connect to MySQL database and insert data into it.
13. Given an objective function $f(x) = x^3 + 3x^2 + 2$, write a Python program that finds the best solution using any heuristic search technique.
14. With an example, explain the process of one-hot encoding.
15. With an example for each, explain the various techniques used in multivariate statistics.
16. Explain how linear regression works with a dataset with two dimensions.
17. Define generators and generator expressions. Explain with an example for each.
18. With suitable code, explain any two preprocessing techniques applied to text data.

(2 x 6 = 12 weight)

PART C
Answer any 2 Questions

19. Prepare detailed note on plotting line plot using `seaborn`. Include code for setting the axis limit, setting a theme, changing the figure size, and scaling the plots.
20. With a computing example, explain the process of clustering along with python code training and testing the model.
21. Prepare a detailed note on the following concepts associated with NumPy module:
    (a). Array creation routines (b). Array Indexing and Slicing
22. Define data serialization. Explain the process of serialization using the required python modules.

(5 x 2 = 10 Weight)