

Roll No.

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

Total No. of Pages : 02

Total No. of Questions : 09

MCA (Sem.-4)

MACHINE LEARNING AND DATA ANALYTICS USING PYTHON

Subject Code : PGCA-1976

M.Code : 91855

Date of Examination : 06-05-2025

Time : 3 Hrs.

Max. Marks : 70

INSTRUCTIONS TO CANDIDATES :

1. SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.
2. SECTION - B & C have FOUR questions each.
3. Attempt any FIVE questions from SECTION B & C carrying TEN marks each.
4. Select atleast TWO questions from SECTION - B & C.

SECTION-A

1. Write short notes on :

- a. What is the cost function used in logistic regression? Discuss how it is optimized.
- b. Describe the process of selecting the number of clusters in K nearest neighbours algorithm.
- c. Define any two criteria used for node splitting in decision tree.
- d. Briefly discuss the principle of classification using Naïve Bayes.
- e. Differentiate between classification and clustering in machine learning.
- f. Write a brief note on NumPy arrays.
- g. What is the difference between a list and a tuple in Python?
- h. How do you create a pie chart using Matplotlib?
- i. What is the purpose of plt.grid() in Matplotlib?
- j. Demonstrate how to handle missing data in DataFrames.

SECTION - B

2. What is a well posed learning problem in machine learning? Give examples. Discuss the steps of designing a machine learning system. Explain the three main types of machine learning.
3. Describe the mathematical principles behind linear regression. Explain how it differs from logistic regression and when each is appropriate.
4. Describe the Random Forest algorithm, including how it works, what ensemble learning is and its advantages over a single decision tree.
5. Discuss the need for dimensionality reduction. Describe how Principal Component Analysis is used for this task. List the characteristics of principal components.

SECTION - C

6. Explain Python's core data types. Provide examples of how to use control structures and define functions.
7. Explain the concepts of Pandas Series and DataFrames. Demonstrate how to create, manipulate and analyze data using these data structures.
8. Explain the use of the groupby() function in Pandas. Demonstrate how to read and write data from/to CSV and Excel files.
9. Demonstrate how to create line plots, scatter plots and bar charts using Matplotlib. Explain how to customize plots with labels, titles and legends.

NOTE : Disclosure of Identity by writing Mobile No. or Marking of passing request on any paper of Answer Sheet will lead to UMC against the Student.