

Roll No.

Total No. of Pages : 03

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 : 07-05-2024

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. Answer Briefly :

- a) When to apply linear regression?
- b) Write the loss function for multiple linear regression.
- c) How to do multiclass classification using logistic regression?
- d) Name a method used in choosing the value of K in KNN.
- e) What are the criteria for choosing a root in a decision tree?
- f) Why do we use the dot (.) operator in python?
- g) What are various datatypes in Python?
- h) When do we need NumPy?
- i) Write two advantages of using PCA.
- j) What is supervised learning? Explain with an example.

SECTION-B

2. a) Explain the working of Decision Trees.
b) What are the advantages of random forests over decision trees?
3. a) Explain PCA in detail using eigen values and eigen vectors.
b) Logistic regression does work of classification. Discuss.
4. a) What are neural networks? When do we use them?
b) What are the various hyperparameters needed for neural networks?
5. Consider the following training set that classifies the output variable play tennis as Yes or No depending upon weather conditions such as Outlook, Temperature, Humidity, and Windy Status.

S. No.	Outlook	Temperature	Humidity	Windy	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rainy	Mild	High	Weak	Yes
5	Rainy	Cool	Normal	Weak	Yes
6	Rainy	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rainy	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rainy	Mild	High	Strong	No

Using Naive Bayes Classifier, classify whether, on a Sunny, Cool, Normal Humidity, and Weak Windy day, we can play tennis or not.

SECTION-C

6. Write Python code for printing the first N Fibonacci series values.
7.
 - a) What are the various functionalities of Pandas library for handling CSV files and manipulating them? Discuss in detail.
 - b) What is the difference between a series and a list? Explain using examples.
8.
 - a) Explain the usage of the Matplotlib library using examples.
 - b) Write the matplotlib command for generating histograms and pie charts.
9.
 - a) What are the various techniques for handling missing data?
 - b) Discuss how to do data input and output in Pandas.

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.