

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MCA DEGREE EXAMINATION MAY 2024
(Third Semester)

Branch – COMPUTER APPLICATIONS

PYTHON FOR MACHINE LEARNING

Time: Three Hours

Maximum: 50 Marks

SECTION-A (5 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (5 x 1 = 5)

- 1 In the data science workflow which stage involves preparing and cleaning the data for analysis?
(i) Acquire Data (ii) Modeling and Evaluation
(iii) Data Munging (iv) Identify the Question
- 2 Which Python library is commonly used for numerical computing and linear algebra operations?
(i) NumPy (ii) Pandas
(iii) Tensorflow (iv) Matplotlib
- 3 Which of the following is the drawback for cordless telephones?
(i) Security (ii) Wireless technology
(iii) Mobile (iv) Limited coverage area
- 4 Wireless Local Area Network is based on _____ standards of IEEE.
(i) 802.10 (ii) 802.11
(iii) 802.12 (iv) 802.13
- 5 IEEE standard for Bluetooth is _____.
(i) 802.15.1 (ii) 802.15.2
(iii) 802.15.3 (iv) 802.15.4

SECTION - B (15 Marks)

Answer ALL Questions

ALL Questions Carry EQUAL Marks (5 x 3 = 15)

- 6 a Outline the key steps in the data science workflow.
OR
b Give an example of how data science is applied in the transportation industry?
- 7 a How do you read data from an external CSV file into a Pandas DataFrame?
OR
b Illustrate the concept of subplots in Matplotlib with example.
- 8 a Analyse the concept of pattern recognition in the context of machine learning.
OR
b What is the curse of dimensionality and why is it a challenge in machine learning?
- 9 a Explain the purpose of ROC CURVES and AUC in binary classification.
OR
b How to improve the performance of a logistic regression model?
- 10 a Name any two common ensemble techniques used in machine learning.
OR
b Distinguish between Bagging and Boosting in Ensemble Learning.

Cont...

SECTION -C (30 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks

(5 x 6 = 30)

- 11 a Explain the roles and responsibilities of data engineers in a data science team.
OR
b Identify the different tools used in data science. Explain in detail.
- 12 a Discuss the following: Lists, Tuples and Dictionaries.
OR
b Enumerate on Indexing and Slicing in Python.
- 13 a Describe some consequences of the curse of dimensionality on data analysis and modeling.
OR
b Illustrate the working concept of K-fold cross-validation and its advantages.
- 14 a Compare and contrast K-means clustering and hierarchical clustering algorithm.
OR
b Evaluate the trade-offs between precision and recall using a confusion matrix.
- 15 a Apply the kernel trick to a simple dataset and explain how it transforms the data for SVM classification.
OR
b Develop a novel deep learning architecture for a specific computer vision task.

Z-Z-Z

END