

B.Sc. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2023.

Fourth Semester

Computer Science – Allied

MACHINE LEARNING TECHNIQUES

(For those who joined in July 2021–2022)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL the questions.

Choose the correct answer:

1. Different learning methods does not include _____
 (a) Memorization
 (b) Analogy
 (c) Deduction
 (d) Introduction

2. Python Programming Language was developed by _____
 (a) Wick van Rossum
 (b) Rasmus Lerdorf
 (c) Guido van Rossum
 (d) Nine Stom
3. In syntax of linear model lm (formula, data, ...), data refers to _____
 (a) Matrix (b) Vector
 (c) Array (d) List
4. The output of training process in machine learning _____
 (a) machine learning algorithm
 (b) machine learning model
 (c) null
 (d) accuracy
5. In dataset loading, the variables of data are called its?
 (a) Response (b) Features
 (c) Target (d) Vector

6. Scikit-learn provides _____ classes.
 (a) 2 (b) 3
 (c) 4 (d) 5
7. _____ is the task of dividing the population or data points into a number of groups.
 (a) Unsupervised learning
 (b) Clustering
 (c) Semi supervised
 (d) Classification
8. Which of the following is not clustering method?
 (a) Density-Based
 (b) Hierarchical Based
 (c) Grid-Based
 (d) Project Based
9. Typical text mining tasks include _____
 (a) Text categorization
 (b) Text clustering
 (c) Entity relation modeling
 (d) All of the above

10. Which of the following technique is not a part of flexible text matching?
 (a) Soundex
 (b) Metaphone
 (c) Keyword Hashing
 (d) Edit Distance

PART B — (5 × 5 = 25 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 250 words.

11. (a) Write notes on Machine Learning.
 Or
 (b) Why Python is suitable for Machine Learning? – Explain.
12. (a) List out the types of Linear Regression.
 Or
 (b) What is the Classification Algorithm? – Explain.
13. (a) Discuss about Probit Regression.
 Or
 (b) Write short notes on Gradient R Algorithm.

14. (a) List out the types of Clustering Methods.

Or

(b) Write short notes on Hierarchical Clustering.

15. (a) What is Vectorizer? – Explain.

Or

(b) Mention the uses of Association Rules.

PART C — (5 × 8 = 40 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 600 words.

16. (a) Describe the classification of Machine Learning.

Or

(b) Write detail note on Python Pandas Data Frame.

17. (a) How to build a Simple Linear Regression Model? – Explain.

Or

(b) Discuss about Decision Tree Classification.

Page 5 Code No. : 20492 E

18. (a) Discuss about Naive Bayes Classifier Algorithm.

Or

(b) Describe any two Advanced Regression Models.

19. (a) Write detail notes on Association Rules.

Or

(b) Explain – How to create Product Segmentation using Clustering?

20. (a) Explain the Naive Bayes Model for sentiment classification.

Or

(b) Write detail notes on K Means Clustering.

Page 6 Code No. : 20492 E

