

B.Tech. DEGREE EXAMINATION, MAY 2023

Fourth Semester

18ECE271T – INTRODUCTION TO DATA SCIENCE

(For the candidates admitted from the academic year 2018-2019 to 2021-2022)

Note:

- (i) **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.
- (ii) **Part - B & Part - C** should be answered in answer booklet.

Time: 3 hours

Max. Marks: 100

PART – A (20 × 1 = 20 Marks)

Marks BL CO PO

Answer ALL Questions

- Computational thinking is the ability to _____ the problem when attacking a large complex task
(A) Select (B) Completion
(C) Merge (D) Decompose
- In Computational thinking problem formulation is also known as
(A) Abstraction (B) Automation
(C) Analyses (D) Completion
- Data Literacy is the ability to extract meaningful information from a _____
(A) Data Controller (B) Library
(C) Dataset (D) Python
- Which of the following instruction need to use in Python to display "Hello. World"?
(A) % java helloworld.java (B) Print ("Hello.world")
(C) Display ("Hello.world") (D) Both (A) and (B)
- Pointiness of data is called as
(A) Skew (B) Kurtosis
(C) Kurtosis (D) Skow
- Lack of symmetry in the data distribution is called as
(A) Kurtosis (B) Kurtosis
(C) Skew (D) Skow
- Removing noise from the data is called as
(A) Normalization (B) Restoration
(C) Smoothing (D) Generalization
- The median is the _____ score of data set
(A) Medium (B) Side
(C) Middle (D) Normal

9. In machine learning , a target is denoted as _____
 (A) Variable (B) Column
 (C) Label (D) Size
10. A variable in statistics is called as _____ in machine learning.
 (A) Label (B) Feature
 (C) Creation (D) Values
11. When do we know the labels on the training examples?
 (A) Supervised Learning (B) Un Supervised Learning
 (C) Both (A) and (B) (D) Reinforcement Learning
12. Which of the following is used to compare the two variables?
 (A) Regression (B) Mugression
 (C) Munging (D) Integration
13. A good method for collecting data is _____
 (A) Discussion (B) Surveys
 (C) Data Collection (D) Importing
14. In survey question types some of the basic Demo-Graphic question are often _____
 (A) DP (B) Describe type Questions
 (C) MCQ (D) Group Questions
15. Interview and focus groups can deliver _____ information.
 (A) Bad, Targeted (B) Rich, Targeted
 (C) Rich (D) Targeted
16. In a survey, information are gathered from _____
 (A) Respondents (B) Informants
 (C) Both (A) and (B) (D) Opposite
17. Qualitative methods implies emphasis on the _____ of entities and on process.
 (A) Qualities (B) Quantities
 (C) Both (A) and (B) (D) Normality
18. BIC is abbreviated as _____
 (A) Bayesian Insight Criterion (B) Basic Information Criterion (BIC)
 (C) Bayesian Information Criterion (BIC) (D) Bayesian Information Case (BIC)
19. The test sample is smaller dataset compared to the _____
 (A) Training Set (B) Model Set
 (C) Sample Set (D) Model Sample Set
20. Cross – Validation also called as _____
 (A) Rotation Estimation (B) Testing
 (C) Circular Estimation (D) Validation

PART – B (5 × 4 = 20 Marks)

Answer ANY FIVE Questions

- | Q. No. | Marks | BL | CO | PO |
|--|-------|----|-----|----------|
| 21. What is the relation between information and data? | 4 | 2 | 1 | 1,1 2 |
| 22. Describe about Data Science. | 4 | 2 | 1 | 1,1 2 |
| 23. Explain about Data Pre-Processing. | 4 | 2 | 2 | 1,1 2 |
| 24. Write about the Data Integration. | 4 | 3 | 2 | 1,1 2 |
| 25. Describe about Machine Learning in short. | 4 | 2 | 3 | 1,5 |
| 26. Summarize about question types. | 4 | 3 | 4,5 | 1,5 |
| 27. What is mean Qualitative Method? | 4 | 2 | 6 | 1,5 |

PART – C (5 × 12 = 60 Marks)

Answer ALL Questions

- | Q. No. | Marks | BL | CO | PO |
|--|-------|----|-----|----------|
| 28. a. Illustrate the computational thinking and tools for Data Science and in detail. | 12 | 2 | 1 | 1,1 2 |
| (OR) | | | | |
| b. Demonstrate how Data Science is used in different fields? | 12 | 2 | 1 | 1,1 2 |
| 29. a. Explain about the Data Pre-Processing in detail with necessary examples. | 12 | 3 | 2 | 1,1 2 |
| (OR) | | | | |
| b. Describe about the following in detail with example. | | 3 | 2 | 1,5 |
| i. Descriptive Analytics | 6 | | | |
| ii. Diagnostics Analytics | 6 | | | |
| 30. a. Write about the decision tree in detail. | 12 | 3 | 3 | 1,5 |
| (OR) | | | | |
| b. Explain about the Support Vector Machine (SVM) with necessary illustration. | 12 | 2 | 3 | 1,5 |
| 31. a. Describe about the different types of survey questions and survey audience in detail. | 12 | 2 | 4,5 | 1,5 |
| (OR) | | | | |
| b. Write about the following in detail with example. | | 3 | 4,5 | 1,5 |
| i. Log and Dairy Data | 6 | | | |
| ii. User Studies in Lab and Field | 6 | | | |
| 32. a. Describe about the following in detail. | | 2 | 6 | 1,5 |
| i. Quantitative methods | 6 | | | |
| ii. Testing | 6 | | | |
| (OR) | | | | |
| b. Explain about the comparing models in detail. | 12 | 3 | 6 | 1,5 |

* * * * *