



Mallikarjuna

Avinashilingam Institute for Home Science and Higher Education for Women
(Deemed to be University Estd, u/s 3 of UGC Act 1956, Category 'A' by MHRD).
Re-accredited with 'A++' Grade by NAAC, CGPA 3.65/4, Category 1 by UGC,
Coimbatore-641043, Tamil Nadu, India

Continuous Internal Assessment Test I – August 2024
SEMESTER-V

Class: III B.Sc

Major: Computer Science

Time: 2 hours

Max. Marks: 60

21BCSC21 Fundamentals of Data Science

Course Outcomes:

At the end of the course, students will

1. Students will be able to apply the basic Data science knowledge on the day problems they encounter.
2. Students will realize that there are various phases that contribute to the completion of a Data Science project and can select among the various modelling techniques.
3. Students will be able to apply Regression techniques for modelling a Data science project.
4. Students will be able to apply the Clustering and Association rule mining for modelling a Data science project.
5. Students can reproduce the knowledge gained and come with a sample case study which they come across in their daily life and implement, document and present the same using the R Tool.

Part-A

6 × 1 = 6

Choose the correct answer

1. What is the most suitable language for Data science? CO1K1
a. Java b. Ruby c. R d. PHP
2. Which of the following graph can be used for simple summarization of data? CO1K2
a. Scatterplot b. overlaying c. Bar plot d. Bar chart.
3. Which of the following is performed by Data Scientists? CO1K1
a. Define the question b. create reproducible code
c. challenge results d. All of the above mentioned
4. Which of the following is not a part of the data science process? CO1K1
a. Communication building b. Operationalize c. Model planning d. Discovery
5. Identify the method used to find relationship between two variables? CO3K1
a. Linear regression b. Logistic Regression c. Multi-calculus d. Both b&c
6. _____ is the method by which you integrate a machine learning model into an existing production environment to make practical business decisions? CO1K1
a. Deployment b. Development c. Documentation d. Maintenance

Part-B

3 × 6 = 18

Answer All the questions

Each answer should not exceed 400 words or four pages

7. a. Explain the importance of Data collection & Management. CO1K1
(or)
b. Explain the steps for Loading Data into R. CO1K2
8. a. Write the steps in Mapping problems to machine learning tasks. CO2K2
(or)
b. Explain the different techniques used in cleaning data. CO1K3
9. a. Highlight the steps in Evaluating scoring models & Evaluating classification Models. CO2K2
(or)
b. Explain the method of Evaluating probability models & Quantifying model soundness. CO2K2

Part-C

3 × 12 = 36

Answer All the questions

Each answer should not exceed 800 words or four pages

10. a. Explain the methods of using graphics & visualization into R. CO1K2
(or)
b. What are the process involved in Deployment & Maintenance. CO1K3
11. a. Illustrate the Ranking model & Clustering model. CO2K3
(or)
b. Illustrate the stages of Data science process with a neat diagram. CO1K3
12. a. Explain the need for Linear Regression model & its applications. CO3K2
(or)
b. Explain the need for Logistic Regression model & its applications. CO3K2
