



Date:

Registration number:

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BENGALURU -27
BCA (DATA ANALYTICS) – VI SEMESTER
SEMESTER EXAMINATION: APRIL 2024
(Examination conducted in May/June 2024)
BCADA6223: DATA MINING AND ANALYTICS
(For current batch students only)

Time: 2 Hours

Max Marks: 60

This paper contains ONE printed page and THREE parts

PART A

Answer ALL the questions

5 x 2 = 10

1. What do you mean by OLAP?
2. Why do we need data integration?
3. What is Data Mart?
4. Describe Descriptive market basket analysis.
5. What do you mean by retail analytics?

PART B

Answer any FIVE questions

5 x 4 = 20

6. Explain any four different types of data that can be mined.
7. What are the stages of data science pipeline? Explain briefly with a neat diagram.
8. When do we use Cosine similarity? Give the advantages of Cosine distance over Euclidian distance.
9. What is Data Warehouse? Explain the three layer architecture of Data Warehouse.
10. Compare and contrast the Data Lake and Data mart with Data warehouse.
11. How do we apply association rule in the field of Healthcare and entertainment? Explain with appropriate examples.
12. How will you apply data analytics in BFSI Sector?

PART C

Answer any THREE questions

3 x 10 = 30

13. Compare and contrast OLAP and OLTP.
14. A retail company is planning to improve data management and analyze sales data from various store locations by keeping all the data in warehouse. The data could be collated to provide a complete overview of the company's operations, enabling them to reallocate resources to expand the business. In this scenario, explain the ETL process with a neat diagram. What are the challenges can be faced during ETL process? (4+2+4=10)
15. How can we find the similarities between two texts of 20 lines each? Explain the details of preprocessing and methodology used. (5+5=10)
16. Explain market basket analysis using Apriori Algorithm in the field of manufacturing.