

Roll No.	1	2	2	0	2	5	8	0	8	2
----------	---	---	---	---	---	---	---	---	---	---

BCA Second Sessional Tests 2023-24

Fourth Semester

BCA4404: Data Warehousing & Data Mining

Time: 3 Hours

Max. Marks: 60

(Section A)

Q.1: Attempt all Questions:

(8\*1=8)

- Define Data mining
- What is KDD?
- What do you understand by pattern?
- Define the concept of Support and Confidence
- What is Regression?
- What is data redundancy?
- Why we need to Pre-process the data?
- Define metadata?

(Section B)

Q2. Attempt any Two:

(6\*2=12)

- Differences between operational Database systems and Data Warehouses?
- What is Data reduction? Discuss in detail?
- Discuss the following
  - Star schema
  - Snow Flake schema
- What do you understand by Metadata? Discuss its types in detail.

(Section C)

Q3. Attempt any Two:

(5\*2=10)

- Write in brief about Data cleaning?
- Compare OLTP and OLAP?
- Explain the Apriori algorithm for finding frequent item sets with an example

Q4. Attempt any Two

(5\*2=10)

- Write short note on nearest neighbor classifier.
- Write short note on:
  - Data transformation
  - Cross-validation.
- Describe Major issues in Data mining?

Q5. Attempt any Two:

(5\*2=10)

- a. What is meant by Market basket analysis?
- b. What is the difference between "supervised" and unsupervised" Learning scheme
- c. What is a "decision tree"?

Q6. Attempt any Two:

(5\*2=10)

- a. What do you meant by Bayesian Classification?
- b. Differentiate between data mining and data warehousing
- c. Explain about the Three-tier data warehouse architecture with a neat diagram.