



44732/E0220

Reg. No.

--	--	--	--	--	--	--	--

V Semester B.C.A. 5 (CBCS) Degree Examination, March/April - 2023

DATA WAREHOUSING AND MINING

(Regular)

Time : 3 Hours

Maximum Marks : 80

Instructions to Candidates :

- 1) All **three** sections are compulsory.
- 2) Draw diagrams wherever necessary.

Section - A

Answer any **ten** of the following.

(10×2=20)

1.
 - a) What is data?
 - b) Define data warehouse.
 - c) What schemas are used in OLAP?
 - d) What is data Mining?
 - e) What is data cleaning?
 - f) Define classification.
 - g) What is OLTP?
 - h) What are frequent patterns?
 - i) What is data mart?
 - j) What is clustering?
 - k) What is partitioning?
 - l) What is weka tool?

Section - B

Answer any **four** of the following questions.

(4×5=20)

2. Enlist any five operation w.r.t OLAP and state their purpose.
3. Describe star schema w.r.t. data warehouse.
4. Explain different steps (stages) in knowledge discovery process.

[P.T.O.]



(2)

44732/E0220

5. List the requirements for cluster analysis and explain any two of them.
6. Draw a neat diagram and explain concept of decision tree.
7. Explain briefly importance of weka tool.

Section - C

Answer any **four** of the following questions.

(4×10=40)

8. Distinguish between OLTP and OLAP system. (10)
9. a) Describe the major tasks involved in data preprocessing. (5)
b) Explain three-tier data warehousing architecture. Draw diagram. (5)
10. a) Briefly explain market basket analysis. (5)
b) Explain Apriori algorithm. (5)
11. a) What is sampling. Explain k-fold cross validation. (5)
b) Explain the terms True Positive(TP), True Negative(TN), False positives (FP) and False Negative (FN) w.r.t. model evaluation. (5)
12. Write note on (any two)
 - a) Iris plant data base.
 - b) Goals of data warehouse.
 - c) Bayesian classification. (5+5)