

**III B. Tech I Semester Supplementary Examinations, June/July-2022**  
**DATA WAREHOUSING AND DATA MINING**

(Computer Science and Engineering)

Time: 3 hours

Max. Marks: 75

Answer any **FIVE** Questions **ONE** Question from **Each unit**

All Questions Carry Equal Marks

\*\*\*\*\*

**UNIT-I**

1. a) Explain data warehouse architecture and models. [8M]  
b) Explain the role of concept hierarchies. [7M]

**(OR)**

2. a) Illustrate OLAP operations. [8M]  
b) Design a data warehouse for a regional weather bureau. The weather bureau has about 1000 probes, which are scattered throughout various land and ocean locations in the region to collect basic weather data, including air pressure, temperature, and precipitation at each hour. All data are sent to the central station, which has collected such data for more than 10 years. Your design should facilitate efficient querying and online analytical processing. [7M]

**UNIT-II**

3. a) Summarize the kinds of patterns that can be mined. [8M]  
b) Explain data transformation strategies. [7M]

**(OR)**

4. a) Briefly outline how to compute the dissimilarity between objects described by the following: [8M]  
i) Normal attributes  
ii) Asymmetric binary attributes  
iii) Numeric attributes  
iv) Term-Frequency vectors  
b) Explain how to handle noisy data. [7M]

**UNIT-III**

5. a) Explain FP-growth algorithm for mining frequent item sets without candidate generation. [8M]

- b) A database has five transactions. Find all frequent item sets using Apriori algorithm with minimum support of 20% and minimum confidence of 80% [7M]

Transaction ID	Items
100	Bread, Cheese, Eggs, Juice
200	Bread, Cheese, Juice
300	Bread, Milk, Yogurt
400	Bread, Juice, Milk
500	Cheese, Juice, Milk

(OR)

6. a) Explain how to mine multilevel association rules. [8M]  
b) Explain pattern evaluation measures. [7M]

UNIT-IV

7. a) Explain how to train a neural network using back propagation technique. [8M]  
b) Explain attribute selection measures with examples. [7M]

(OR)

8. a) Illustrate Naïve Bayesian classification. [8M]  
b) Elaborate on various measures that are appropriate to evaluate the performance of a classification model. [7M]

UNIT-V

9. a) Summarize basic clustering methods. [8M]  
b) Explain how to cluster high dimensional database. [7M]

(OR)

10. a) Illustrate k-means clustering algorithm. [8M]  
b) Distinguish Agglomerative hierarchical clustering from divisive hierarchical clustering. [7M]

\*\*\*\*\*