



**SIDDHARTH GROUP OF INSTITUTIONS :: PUTTUR**  
Siddharth Nagar, Narayanavanam Road – 517583

**QUESTION BANK (DESCRIPTIVE)**

**Subject with Code :** Data Warehousing and Mining (16MC815)      **Course & Branch:** MCA  
**Year & Sem:** II-MCA & I-Sem      **Regulation:** R16

**Unit 1**

1. Describe data mining. In your answer, address the following:
  - a) Is it another type? 2m
  - b) Is it a simple transformation of technology developed from databases, statistics, and machine learning? 2m
  - c) Explain how the evolutions of data base technology lead to data mining. 4m
  - d) Describe the steps involved in data mining when viewed as a process of knowledge discovery 4m
2. Distinguish between the data warehouse and databases. How they are similar? 12m
3. Distinguish between the data warehouses and data mining. 12m
4. a) Explain the difference between discrimination and classification 4m  
b) Between characterization and clustering 4m  
c) Between classification and prediction 4m
5. a) Discuss briefly about the data smoothing techniques 6m  
b) Differentiate operational database systems and data warehousing 6m
6. a) Explain data mining as a step in the process of knowledge discovery. 6m  
b) Describe briefly the concept hierarchy generation for numerical data? 6m
7. a) Discuss about the concept hierarchy generation for categorical data? 6m  
b) Describe the various data reduction techniques? 6m
8. Define data cleaning. Express the different techniques for handling missing values. 6m
9. a) Discuss issues to consider during data integration 6m  
b) Explain about the various data smoothing techniques 6m

10. List and describe the five primitives for specifying a data mining task 12m

### Unit 2

1. a) Differentiate operational database systems and data warehousing. 6m  
b) Discuss briefly about the multidimensional data models. 6m
2. a) Explain with an example the different schemas for multidimensional databases? 6m  
b) Demonstrate the four major types of concept hierarchies are schema hierarchies, set-grouping hierarchies, operation-derived hierarchies and rule-based hierarchies. Briefly define each type of hierarchy. 6m
3. Describe the three-tier data warehousing architecture 12m
4. a) Discuss the efficient processing of OLAP queries. 6m  
b) Explain the data warehouse applications. 6m
5. a) Explain the architecture for on-line analytical mining. 6m  
b) Illustrate the applications of data mining. 6m
6. a) Explain the efficient methods for data cube computation. 6m  
b) Describe the common techniques are used in ROLAP and MOLAP. 6m
7. a) Explain how to compute iceberg cubes by using BUC and star-cubing algorithms. 6m  
b) Discuss the complex iceberg condition to compute cube. 6m
8. Explain about the concept description? And what are the differences between concept description in large databases and OLAP? 12m
9. a) State and explain algorithm for attribute-oriented induction 6m  
b) Describe mining class comparisons and class description. 6m
10. a) Compare the schemas for the multidimensional data models. 6m  
b) Explain about the data warehouse implementation with an example. 6m

### UNIT 3

1. a) Define the terms frequent itemsets, closed itemsets and association rules. 6m

- b) Describe the different techniques to improve the efficiency of Apriori? Explain. 6m
2. Discuss the FP-growth algorithm. Explain with an example. 12m
3. a) Discuss about mining multilevel association rules from transaction databases in detail. 6m
- b) Explain how to mine the multidimensional association rules from relational databases and data warehouses 6m
4. a) Discuss about constraint-based association mining. 6m
- b) Discuss the generating association rules from frequent itemsets. 6m
5. Explain the Apriori algorithm with given example. 12m

TiD	List of item ids
T100	1,2,5
T101	2,4
T103	2,3
T104	1,2,4
T105	1,3
T106	2,3
T107	1,3
T108	1,2,3,5
T109	1,2,3

6. Explain about the classification and prediction. Example with an example. 12m
7. Discuss about basic decision tree induction algorithm. 12m
8. Discuss the back propagation algorithm and explain with example. 12m
9. a) Explain briefly various measures associated with attribute selection. 6m
- b) Explain training of Bayesian belief networks. 6m
10. a) Explain about “IF\_THEN” rules used for classification with an example. 6m
- b) Discuss the process of extracting IF-THEN rules using sequential covering algorithm. 6m

### UNIT IV

- |     |  |     |
|-----|--|-----|
| 1.  | Discuss the various types of data in cluster analysis.                                   | 12m |
| 2.  | Explain the categories of major clustering methods.                                      | 12m |
| 3.  | a) Write algorithms for k-means and k-medoids. Explain.                                  | 6m  |
|     | b) Describe the different types of hierarchical methods.                                 | 6m  |
| 4.  | Demonstrate about the following hierarchical methods a)BIRCH b)Chamelon.                 | 12m |
| 5.  | a) Discuss about the DBSCAN density-based methods.                                       | 6m  |
|     | b) Explain about grid-based methods.   | 6m  |
| 6.  | a) Describe the mode-based methods.  | 6m  |
|     | b) Explain the working of CLIQUE algorithm.  | 6m  |
| 7.  | Explain about the mining of data streams.  | 6m  |
| 8.  | Discuss the four major components of trend analysis for characterizing time series data? | 12m |
| 9.  | a) Demonstrate about the similarity search in time series analysis.                      | 6m  |
|     | b) Describe about the sequential pattern mining.   | 6m  |
| 10. | a)Describe the characteristics of social networks.                                       | 6m  |
|     | b)List the tasks and challenges of link mining.  | 6m  |

### UNIT V

- |    |  |     |
|----|--|-----|
| 1. | Summarize the descriptive mining of complex data objects.      | 12m |
| 2. | a)Discuss briefly about the generalization of structured data. | 6m  |
|    | b) Define class composition hierarchy.                         | 6m  |
| 3. | Explain how it is generalized by giving a suitable example.    | 12m |
| 4. | a) Explain the construction and mining of object cubes.        | 6m  |

- b) Describe the generalization-based mining of plan databases by divide-and-conquer with an example. 6m
5. Explain the construction of spatial data cube with suitable example. 6m
6. a) Describe multimedia databases. 6m  
b) Explain mining multimedia databases. 6m
7. a) Explain briefly about the text data analysis and information retrieval. 6m  
b) Describe about the Latent Semantic Indexing (LSI). 6m
8. a) Discuss about the Probabilistic Latent Semantic Indexing (PLSI). 6m  
b) Explain about the Locality Preserving Indexing (LPI). 6m
9. a) Explain about mining the world wide web. 6m  
b) Describe about the mining the web page layout structure. 6m
10. a) Demonstrate about the data mining for intrusion detection. 6m  
b) Explain the examples of commercial data mining systems. 6m

*Prepared by Dr. SABEEN S, Professor, Department of MCA*