

IV B.Tech II Semester Regular Examinations, Apr/May 2006
DATA MINING AND DATA WARE HOUSING
(Computer Science & Systems Engineering)

Time: 3 hours

Max Marks: 80

Answer any FIVE Questions
All Questions carry equal marks

1. (a) Discuss Business Case analysis, Education and Prototyping.
(b) Explain Business requirements analysis and technical blueprint.
(c) Describe the requirements Evolution and History load. [6+6+4]
2. Describe the following options while designing FACT table:
(a) Identifying significant historical period.
(b) Sampling for detailed data.
(c) Column selection. [6+4+6]
3. Describe the role and importance of hardware architecture of a data warehouse. Explain with the help of an example situation. [16]
4. (a) Explain the need and role of security on the performance of data warehouse
(b) Describe the impact of security on the design of the data warehouse. [8+8]
5. Estimate the Disk space required for a data warehouse. [16]
6. (a) Explain in detail about the "COST of Encoding TREE".
(b) Explain about the PRUNING algorithm. [8+8]
7. (a) What is the underlying principles of "The Hidden Web"? How is text mining related to web mining? What are the techniques of text mining? [3+2+3]
(b) What is Page Rank? How is it computed? [3+5]
8. (a) Why is concept hierarchy important for spatial data? Identify certain cases of occurrences of such hierarchies in spatial data.
(b) How do you handle spatial or temporal data for decision tree construction? [8+8]

IV B.Tech II Semester Regular Examinations, Apr/May 2006
DATA MINING AND DATA WARE HOUSING
(Computer Science & Systems Engineering)

Time: 3 hours

Max Marks: 80

Answer any FIVE Questions
All Questions carry equal marks

1. (a) How to clear and transform the Data?
(b) Explain how to transforming into Effective Structures?
(c) Describe the Backup and Archive process. [6+4+6]
2. (a) When is a summary table too big to be useful ?
(b) Relate and discuss the various degrees of aggregation within summary tables. [8+8]
3. Describe the operational design issues involved in the data warehouse system. Explain with the help of an example situation. [16]
4. Describe the following terms.
(a) Operational system
(b) Mission critical system
(c) 7 x 24 system
(d) 7 x 24 x 52 system. [4+4+4+4]
5. (a) Is daily processing different from overnight processing for Load estimation process?
(b) What are the system administration requirements of database siting. [10+6]
6. What is splitting criteria? With an example explain about the [2]
(a) Class Histogram, and [7]
(b) Count Matrix. [7]
7. (a) What is concept hierarchy? How is it related to web mining? [3+5]
(b) Which frequent itemset mining is suitable for text mining and why. Explain? [8]
8. (a) Discuss the major algorithms of the sequence mining problem. [8]
(b) What is the event-prediction problem? Propose one algorithm to solve this problem. [3+5]

IV B.Tech II Semester Regular Examinations, Apr/May 2006
DATA MINING AND DATA WARE HOUSING
(Computer Science & Systems Engineering)

Time: 3 hours

Max Marks: 80

Answer any FIVE Questions
All Questions carry equal marks

1. (a) Explain with the help of a neat sketch a data warehouse with detailed data in archived storage.
(b) What is Query Manager? How is it different from load manager? [8+8]
2. (a) Explain Views and Synonymms on fact tables.
(b) Explain Multidimensional Schema. [8+8]
3. Compare and contrast the differences between the symmetric multi-processing and massively parallel processing architectures of data warehouse systems. [16]
4. (a) Describe the role of software to implement backup strategy of a data warehouse system.
(b) What are the distinct features of software to implement backup strategy. [8+8]
5. (a) Elaborate estimation of CPU band width for the different phases involved in loading data warehouse.
(b) How do you estimate the memory requirement for a data warehouse. [10+6]
6. What is a DECISION TREE? With an example, Explain about the CART, ID3 algorithms. Give comparison between CART & ID3 algorithm. [3+9+4]
7. (a) What are the different types of web mining? How is web usage mining different from web structure mining and web content mining? [4+4]
(b) What is Page Rank? How is it computed? [3+5]
8. (a) Discuss the major features of the timeweaver algorithm.
(b) Describe the working of the SPADE algorithm. [8+8]

IV B.Tech II Semester Regular Examinations, Apr/May 2006
DATA MINING AND DATA WARE HOUSING
(Computer Science & Systems Engineering)

Time: 3 hours

Max Marks: 80

Answer any FIVE Questions
All Questions carry equal marks

1. (a) What is a Data Warehouse? Discuss in detail.
(b) Describe with the help of a figure the typical process flow within a Data Warehouse. [8+8]
2. (a) Explain Views and Synonymms on fact tables.
(b) Explain Multidimensional Schema. [8+8]
3. Describe the operational design issues involved in the data warehouse system. Explain with the help of an example situation. [16]
4. (a) Explain the need and role of security on the performance of data warehouse
(b) Describe the impact of security on the design of the data warehouse. [8+8]
5. How do you make load estimation for a data warehouse? [16]
6. (a) Describe the estimation of gini index for a node used in the public algorithm.
(b) Describe different pruning strategies. [8+8]
7. (a) What is the underlying principles of “The Hidden Web”? How is text mining related to web mining? What are the techniques of text mining? [3+2+3]
(b) What is Page Rank? How is it computed? [3+5]
8. (a) What is spatial mining? Explain about the spatial mining tasks? [3+5]
(b) Write and explain about spatial trend algorithm. [8]
