

**IV B.Tech I Semester Regular Examinations, November 2005**  
**DATA MINING AND WAREHOUSING**  
**( Common to Computer Science & Engineering and Information Technology)**

**Time: 3 hours**

**Max Marks: 80**

**Answer any FIVE Questions**  
**All Questions carry equal marks**

\*\*\*\*\*

1. Give an introduction of Delivery process in Data Warehouse. [16]
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 requirements with respect to service level agreements.  
(a) User requirements.  
(b) System requirements. [8+8]
5. How much memory is needed and how is it estimated? [16]
6. (a) Write in detail about the Testing the operational environment.  
(b) Write in detail about the Testing of Database Performance. [10+6]
7. (a) What is concept hierarchy? How is it related to web mining? [3+5]  
(b) What is page Rank, and How is it computed? How it is related to web mining? [3+2+3]
8. (a) How do you handle spatial and non-spatial data, while carrying out any mining task?  
(b) Propose different neighborhood relationships that can be used for density-based clustering of spatial data. [8+8]

\*\*\*\*\*

**IV B.Tech I Semester Regular Examinations, November 2005**  
**DATA MINING AND WAREHOUSING**  
**( Common to Computer Science & Engineering and Information Technology)**

**Time: 3 hours****Max Marks: 80**

**Answer any FIVE Questions**  
**All Questions carry equal marks**

\*\*\*\*\*

1. Give an introduction of Delivery process in Data Warehouse. [16]
2. (a) Explain difference between designing a Data Warehouse and an OLTP system.  
(b) Explain fact table identification process. [8+8]
3. Describe the role and importance of hardware architecture of a data warehouse. Explain with the help of an example situation. [16]
4. (a) Why is it important to get all the security and audit requirements clearly documented ?  
(b) “Data movement is an expensive process“ Justify [8+8]
5. (a) What is the significance of performance assessment before tuning the warehouse.  
(b) Explain the AD hoc query tuning mechanism in a data warehouse. [10+6]
6. (a) What is a Decision Tree? What are the advantages and disadvantages of DECISION TREE classifications? [3+5]  
(b) For the given data set create a Decision Tree? And explain about the knowledge obtained from it. [4+4]

OUTLOOK	TEMP(F)	HUMIDITY(%)	WINDY	CLASS
sunny	79	90	True	play
sunny	56	70	Flase	play
sunny	79	75	True	no play
sunny	60	90	True	no play
overcast	88	88	False	no play
overcast	63	75	True	play
overcast	88	95	False	play
Rain	78	60	False	play
Rain	66	70	False	no play
Rain	68	60	True	play

7. (a) What is concept hierarchy? How is it related to web mining? [3+5]  
(b) What is page Rank, and How is it computed? How it is related to web mining? [3+2+3]

8. (a) What is spatial trend? Explain about the spatial trend detection algorithm. [3+5]
- (b) What is spatial clustering? Write about spatial characterization. [3+5]

★ ★ ★ ★ ★

**IV B.Tech I Semester Regular Examinations, November 2005**  
**DATA MINING AND WAREHOUSING**  
**( Common to Computer Science & Engineering and Information Technology)**

**Time: 3 hours**

**Max Marks: 80**

**Answer any FIVE Questions**  
**All Questions carry equal marks**

\*\*\*\*\*

1. (a) Explain metadata and Data Marting concepts briefly with reference to data warehousing.  
(b) Draw the three tire decision support information diagram depicting the summary and detailed information. [8+8]
2. (a) Explain Designing dimension tables.  
(b) Explain Designing Starflake Scheme. [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) What are the different parts of the data warehouse which are affected by the security restrictions.  
(b) Describe the security requirements and access restrictions on the subsets of the data. [8+8]
5. Describe Daily processing with reference to load estimation. [16]
6. (a) Explain about the Three basic levels of Testing.  
(b) Write in detail about the stages in Developing the Test Plan. [8+8]
7. (a) What is the underlying principles of “*TheHiddenWeb*”? How is text mining related to web mining? What are the techniques of text mining?  
(b) Discuss about [8+8]
  - i. Transverse & Intrinsic Links,
  - ii. Reference Nods & Index nodes.
8. (a) Describe different similarity measures of time-series data.  
(b) Discuss the major features of the timeweaver algorithm. [8+8]

\*\*\*\*\*

**IV B.Tech I Semester Regular Examinations, November 2005**  
**DATA MINING AND WAREHOUSING**  
**( Common to Computer Science & Engineering and Information Technology)**

**Time: 3 hours**

**Max Marks: 80**

**Answer any FIVE Questions**  
**All Questions carry equal marks**

\*\*\*\*\*

1. Explain the role of extract and load process with an emphasis on the following topics.
  - (a) Controlling the process
  - (b) Snapshot and initiating the extract.
  - (c) Loading the data. [16]
2.
  - (a) Explain horizontal partitioning.
  - (b) Explain vertical partitioning. [10+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) Describe the role access hierarchy of a data warehouse.
  - (b) The tighter the security the more person oriented. Do you accept the statement? Offer your remarks with justification. [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.
  - (a) Describe the ID3 algorithm of the DECISION TREE construction. Why is it unsuitable for DATA MINING applications? [4+4]
  - (b) Explain about CHAID algorithms. [8]
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) How do you handle spatial and non-spatial data, while carrying out any mining task?
  - (b) Propose different neighborhood relationships that can be used for density-based clustering of spatial data. [8+8]

\*\*\*\*\*