(Please write your Roll No. immediately) Roll No. ……

End-Term Examination

Fourth Semester [MCA] – MAY-JUNE 2006

Paper Code: MCA-202 (ID -44202) Subject: Data Warehousing & Data Mining

Time: 3 Hours **Maximum Marks: 60**

Note: Question no. 1 is compulsory and carries 20 marks. Apart from question 1, the question paper consist of four units each containing two questions carrying 10 marks each. You must attempt one question from each unit.

Q. 1 (a) Define information crisis.
(b) How data is different from information?
(c) Provide functional definition of a data warehouse.
(d) Which type of activities are covered under data cleaning.
(e) Explain briefly the significance of JAD techniques.
(f) How ER modeling is different from dimensional modeling?
(g) A dimension table is wide; the fact table is deep. Do you agree? Why / why not?
(h)Are the junk dimensions necessary in a data warehouse?
(i) What do you understand by slice-and dice.
(j) Is the data warehouses a pre-requisite for data mining. Why/ why not.

UNIT -1

Q. 2 Data warehousing is the only viable means to resolve the information crisis and to provide strategic information. List five reasons to support this assertion and explain them.

Q. 3 How are the top-down and bottom-up approaches for building a data warehouse different? Discuss the advantages and disadvantages of each approach.

UNIT -II

Q. 4 Why is entity-relationship modeling technique not suitable for the data warehouse? How is dimensional modeling different?

Q. 5 How does a snowflake scheme differs from a STAR schema. Explain with an example. Name any two disadvantages of snowflake schema.

UNIT -III

Q. 6 What are the essential differences between MOLAP and ROLAP models. Also list a few similarities.

Q. 7 State any five of Dr. Cod's guidelines for an OLAP system, giving brief description of each.

UNIT -1V

Q. 8 Discuss various application of data mining.

Q. 9 How OLAP is different from data mining. Explain using example.

(Please Write your Exam Roll No. immediately) Roll No. ……..

END-TERM EXAMINATION

FIFTH SEMESTER [MCA] - DECEMBER 2004

Paper Code: MCA-323 **Subject: Data Warehousing and Data Mining**

Time: 3 Hours **Maximum Marks: 60**

Note: Attempt five questions in all including Q. 1 which is compulsory.

Q. 1. Compulsory Question; Attempt any six parts **12**

(a) What are the conceptual features of Data Warehousing which makes it superior to conventional DBMS?
(b) Differentiate data mart and data warehouse
(c) Multi dimensional databases versus two dimensional conventional databases.
(d) Legacy and operational databases
(e) Compare data mining concepts with conventional mining engineer.
(f) Describe ROLAP
(g) Data Integration
(h) Cluster analysis and its application
(i) Discuss contents of meta data repository.

Q. 2. (a) Differentiate operational database system and data warehouses. Some times data warehousing also is termed "from tables and spread sheets to data cubes". Comment **4**

(b) Discuss Data Warehouse architecture, specifically describe 3-tier data warehouse architecture. **8**

Q. 3. (a) Describe various approaches of data mining. **5**

(b) CRM (Customer Relation Management) is considered to be and ideal application of data mining. Illustrate this if you were given data on spending habit of famous credit card holders. **7**

Q. 4. What are the difference between three main types of data warehouse usage, information processing, analytical processing, and data mining? Also discuss the motivation behind OLAP mining (OLAM). **12**

Q. 5. (a) Describe various steps to build a data warehouse in an organization, say in banking sector. **6**

(b) What is level of granularity, how will you decide on the level of granualering in your data for data warehouse. **4**

Q. 6. (a) KDD is really a good example of convergence of technologies where disciplines like statistics, graphics, mathematical and other analytical tools support KDD. Discuss with example. **6**

(b) Data warehousing and data mining concepts have provided new approval for DSS. Discuss. **6**

Q. 7. (a) Differentiate stationary, distributed and virtual data warehouses. **6**

(b) Describe use of Fuzzy logic and its tools in data mining. **6**

Q. 8. Write short notes on any two topics **12**

(a) Data mining using neural versus genetic algorithm
(b) Data web – its applications
(c) Data models on Data warehousing
(d) Data warehousing system for EIS.

(Please Write your Exam Roll No. immediately) Roll No. ……..

END-TERM EXAMINATION

FIFTH SEMESTER [MCA] - DECEMBER 2002

Paper Code: MCA-323 **Subject: Data Warehousing and Data Mining**

Time: 3 Hours **Maximum Marks: 60**

Note: Attempt any five questions.

Q. 1. (a) What can data mining do? What are the various approaches of Data Mining? **6**

(b) What are the goals and components of a data warehouses? Describe its scope and practical implications. **6**

Q. 2. (a) Compare LAN based data warehouse with stage data warehouse. **6**

(b) How do differentiate between stationary, distributed and virtual data warehouses. Give appropriate example. **6**

Q. 3. Differentiate between the following :-

(e) 2-tier, 3-tier and 4-tier data warehouses. **8**

(f) Meta data and operational data. **4**

Q. 4. Discuss knowledge discovery through statistical techniques in detail. Compare this with knowledge discovery through neural networks. **12**

Q. 5. What is a Datamart? What are the advantages of using OLAP databases for decision support? Give appropriate example. **12**

Q. 6. Explain the difference between OLTP and OLAP? What are the various DSS topologies and multidimensional databases? Explain with the help of appropriate examples. **12**

Q. 7. (a) When do we prefer to use Neural Networks in data mining? What are the limitations and consequences of choosing neural networks in DSS? **6**

(b) What is Data partitioning? How do we use data mining for Customer relation Management (CRM)? **6**

Q. 8. Write short notes on any two topics **6+6**

(g) Fuzzy techniques for Data Mining

(h) Data Mining using Genetic Algorithms

(i) Data warehouse architecture
