

3/12/14/FW  
LIB

Reg. No. : 

--	--	--	--	--	--	--	--	--	--	--	--

**Question Paper Code : 45304**

5 Year M.Sc. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2014.

Elective

Software Engineering

ESE 514 — DATA MINING AND DATA WAREHOUSING

(Regulation 2010)

Time : Three hours

Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. What are the core components that form the architecture of a typical data mining system?
2. Differentiate between operational database systems and data warehouses.
3. Define the term "Gini-Index".
4. What is a noisy data? Give the binning methods for data smoothing.
5. Data warehouses must be architected. Why?
6. List out the operations that are supported by the load manager entity of a data warehouse.
7. Give the merits in the use of log files and trace files for a data warehouse.
8. What are the key tasks to be accomplished during overnight processing to ensure reliable business day?
9. Write down the formula used for calculating the space required for storing the database.
10. List down the objective measures of performance that are considered for tuning the data warehouse.

PART B — (5 × 16 = 80 marks)

11. (a) (i) Illustrate with a suitable example about the role of data mining in a marketing scenario.
- (ii) Discuss in detail about the various data mining task primitives.

Or

- (b) (i) Elaborate on the merits and demerits of the information and production factor in data mining.
- (ii) Explain in detail about the self-learning computer systems and its role in decision making.
12. (a) (i) Write the sequential covering algorithm used in association rule based induction.
- (ii) What do you mean by Decision Tree Induction? Explain with an example.

Or

- (b) (i) Explain in detail about the preliminary analysis of the data sets carried out through query tools.
- (ii) Explain in detail about the steps involved in the process of knowledge discovery through data mining.
13. (a) (i) Discuss in detail about how the aggregation of data is carried out in a data warehouse.
- (ii) Enumerate in detail about the process architecture that forms the integral part in a data warehouse.

Or

- (b) (i) Write short notes on the system and data warehouse managers.
- (ii) Discuss in detail about the statement "Data marts should not be used as an alternative to aggregation".
14. (a) (i) Enumerate in detail about the need for parallel index build for enabling index maintenance within a reasonable timescale.
- (ii) Why should the service level agreement of a data warehouse must covers all the dependencies that it has with the outside real world? Explain.

Or

- (b) (i) Discuss in detail about the possible ways of ensuring security for a data warehouse.
- (ii) Elaborate on the hardware and operational design of a data warehouse.

15. (a) (i) Enumerate in detail about the promising features that are considered to be inevitable in any Data warehouse.
- (ii) Compare and contrast fixed queries with Ad hoc queries.

Or

- (b) (i) Write short notes on the steps involved in testing backup recovery.
- (ii) Elaborate on the process of testing database performance with an example.
-