LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – STATISTICS

THIRD SEMESTER – NOVEMBER 2009

ST 3955 - DATA WAREHOUSING AND DATA MINING

Date & Time: 12/11/2009 / 9:00 - 12:00 Dept. No. | | Max. : 100 Marks

SECTION – A

Answer ALL questions. (10 x 2 = 20 marks)

1. Define Data Warehouse.
2. Explain the terms dimension data and fact data.
3. What are the two disadvantages of using OLAP?
4. State the reasons for creating a data mart.
5. What is meant by a business model?
6. Write the advantages of using Bitmap indexes in Data Warehousing.
7. Mention the three main types of parallel processing implemented in ETL applications.
8. Describe the purpose of compressing tables.
9. Define a Cluster. What are the variable types that occur in Cluster Analysis?
10. Explain the term Business Intelligence. What are the common functions of Business Intelligence Technologies?

SECTION – B

Answer any FIVE questions. (5 x 8 = 40 marks)

11. Describe Online Transaction Processing (OLTP) and mention the uses of this technology.
12. Discuss the benefits of using Decision Support Systems (DSS).
13. Explain the purpose of creating a function model and hence describe a functional modeling perspective.
14. What is a Business Model Design? How is it distinct from Business Modelling?
15. Give an example of a snowflake schema used in a query.
16. What are the advantages of partitioning Tables and Indexes?
17. Discuss the preprocessing steps that may be applied to the data to help improve the accuracy, efficiency, and scalability of the classification or prediction process.
18. State the advantages and disadvantages of Multidimensional Online Analytical Processing (MOLAP) technology.

SECTION – C

Answer any TWO questions. (2 x 20 = 40 marks)

19(a). Explain the differences between Data Warehouse and OLTP Systems. (12)

(b). Explain the term Enterprise Architecture in detail. (8)

20(a). Mention the steps to be taken to convert logical design to physical implementation. (12)

(b). Write the execution steps involved in a real – life ETL cycle. (8)

21(a). What is back propagation? How does it work? (8)

(b). Discuss the two – step process of association rule mining and give an example. (12)

22(a). Explain the different approaches used for similarity based retrieval in image databases based on image signature. (12)

(b). Describe the Document Ranking Method for information retrieval. (8)

-------x--------x--------x--------