USN | | | | | | | | | | |

CS845

Eighth Semester B.E. Degree Examination, Dec. 07 / Jan. 08
Data Mining and Warehousing

Time: 3 hrs. Max. Marks:100

Note : ***Answer any FIVE full questions.***

1 a. What is data warehousing? With neat diagram explain the components of data warehouse architecture. **(10 Marks)**

b. With an example, explain all OLAP operations in the multidimensional Data Model. **(10 Marks)**

2 a. Discuss the following terms in brief : i) Data cleaning ii) Data integration and transformation iii) Data reduction. **(10 Marks)**

b. Explain Data Mining Functionalities and classification of Data Mining Systems. **(10 Marks)**

3 a. What do you mean by presentation and visualization in Data Mining? **(04 Marks)**

b. What is an association rule in data mining? Explain various kinds of association rules. **(10 Marks)**

c. A database has six transactions of purchase of books from a bookshop as given. **(06 Marks)**
t_1: = { ANN, CC, TC, CG } , t_2 : = { CC, D, CG } , t_3 : = { ANN, D, CC, TC }
t_4 : = {ANN, CC, D, CG } , t_5: = { ANN, CC, D, TC, CG }, t_6:= { CC, D, TC}
Let X = { CC, TC} and Y = { ANN, TC, CC}.
Find the confidence and support of the association rule $X \Rightarrow Y$ and inverse rule $Y \Rightarrow X$.

4 a. A partitioning variation of Apriori subdivides the transactions of a database D in to 'n' non overlapping partitions. Prove that any item set that is frequent in D must be frequent in at least one partition of D. **(06 Marks)**

b. Define classification and prediction. What are the steps involved in preparing the data for classification and prediction. **(10 Marks)**

c. Write about Decision Tree Induction. **(04 Marks)**

5 a. Explain Bayesian classification. **(08 Marks)**

b. What are different types of data in cluster analysis? Explain each with example. **(12 Marks)**

6 a. Explain the following methods in clustering analysis i) Partitioning ii) Hierarchical. **(10 Marks)**

b. Mention two approaches of model based clustering methods and explain each in detail. **(10 Marks)**

7 a. Explain any three Data Mining Applications. **(12 Marks)**

b. What is the difference between visual data mining and data visualization. **(08 Marks)**

8 Write short notes on any two :

a. Data Mining Query language.

b. Clustering with obstacle objects.

c. Theoretical foundations of Data Mining. **(20 Marks)**

*****