Total No. of Questions : 10]

SEAT No. :

P4008

[Total No. of Pages : 2

[4760] - 412

M.E. (Computer Engineering)

DATA WAREHOUSING AND DATA MINING (Elective - IV)

(2008 Pattern)

Time :3 Hours] *[Max. Marks :100*

Instructions to the candidates:-

1) ***Answers to the two sections should be written in separate books.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Assume suitable data, if necessary.***
4) ***Section I : Q1 is compulsory. Solve Any Two questions out of Q2,Q3,Q4,Q5.***
5) ***Section I I: Q6 is compulsory. Solve Any Two questions out of Q7,Q8,Q9,Q10.***

SECTION - I

Q1) a) Explain Data Warehouse Load Manager (ETL process) in details. **[8]**

b) Design snowflake schemas for financial services data warehouse. **[8]**

Q2) a) Define Multidimensional Model. Explain different OLAP servers. **[9]**

b) Explain various guidelines for designing fact table and dimension table in conceptual modeling schemas. **[8]**

Q3) a) Describe Market Basket problem using Apriori Association Rule mining algorithm. **[9]**

b) Describe data mining primitives with suitable example. **[8]**

Q4) a) Write a short Note on : **[12]**

i) Data Decretization

ii) Materialized View

b) Explain correlation analysis technique with suitable example. **[5]**

P.T.O.

Q5) a) Explain mining Multi-level Association Rules from Relational Database and Data Warehouse. **[9]**

b) Describe dimensionality reduction techniques in data warehouse **[8]**

SECTION - II

Q6) a) Define Classification and Prediction. Explain decision tree based classification method with suitable example. **[8]**

b) Define Information Retrieval System. Describe keyword based association text mining. **[8]**

Q7) a) Describe K-Nearest Neighbor classifiers with suitable example. **[9]**

b) Define Outlier. Explain any two methods for outlier detection and handling. **[8]**

Q8) a) Explain DBSCAN density Clustering algorithm with suitable example. **[9]**

b) Write a short Note on : **[8]**

i) Spatial association

ii) Web usage mining

Q9) a) Explain Naïve Bayesian classification algorithm for email classification. **[9]**

b) Describe K-Means clustering technique with suitable example. **[8]**

Q10) a) Explain neural network approach used for model based clustering. **[9]**

b) Explain Page rank algorithm in web and text mining. **[8]**

XXXX