Total No. of Questions : 10]

SEAT No. : ____

P4902

[Total No. of Pages : 2

[5155]-28

M.E. (Computer Engineering) (Semester - II)

DATA WAREHOUSING AND DATA MINING

(2008 Pattern) (Elective - IV)

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

Instructions to the candidates:

1) ***Q.1 and Q.6 are compulsory.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Assume suitable data, if necessary.***
4) ***Solve any two questions from Q.2,Q.3,Q.4,Q.5.***
5) ***Solve any two questions from Q.7,Q.8,Q.9,Q.10.***
6) ***Answers to the two sections should be written in separate books.***

SECTION - I

Q1) a) Explain design methodology of data warehouse. **[8]**

b) Explain multidimensional model with suitable example. **[8]**

Q2) a) Discuss issues to be considered during data integration. **[8]**

b) Why it is necessary to preprocess data? Explain in detail Data cleaning steps. **[9]**

Q3) a) What is concept description? Explain summarization based characterization. **[8]**

b) What is association mining? Explain Apriori algorithm. **[9]**

Q4) a) Explain different methods for handling data redundancy in data integration. **[8]**

b) Explain frequent pattern growth algorithm with suitable example. **[9]**

P.T.O.

Q5) Write a short note on **[17]**

a) OLAP

b) corelation analysis

c) data mining primitives

SECTION - II

Q6) a) Explain classification using decision trees. **[8]**

b) Explain different parameters to evaluate classification model. **[8]**

Q7) a) Explain the different distance measures in clustering mining task. **[8]**

b) Explain k-means clustering with suitable example. **[9]**

Q8) a) Explain spatial association with suitable example. **[8]**

b) Explain keyword association and document classification method in text mining. **[9]**

Q9) a) Define information retrieval system.Describe vector space model. **[8]**

b) Explain Rough set data mining approach with suitable example. **[9]**

Q10) Write a short note on **[17]**

a) Genetic algorithms

b) web mining

c) outlier analysis