[3764]-416
B.E. (Computer Engineering)
ADVANCED DATABASES
(2003 Course) (410445) (Elective - I)

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) Figures to the right indicate full marks.
- 4) Assume suitable data, if necessary.

SECTION - I

Q1) a) Explain any two partitioning techniques with respect to parallel database system in detail. **[8]**

b) Explain the following with respect to parallel database system.

i) Intraoperation and interoperation parallelism. **[8]**

ii) Fragment - and - Replicate join.

OR

Q2) a) Explain interquery and intraquery parallelism. **[8]**

b) Explain the following with respect to parallel database system.

i) Range – Partitioning sort. **[8]**

ii) Skew.

Q3) a) Describe and compare homogenous and heterogenous databases with respect to distributed databases. **[4]**

b) Explain deadlock handling with respect to distributed databases. **[6]**

c) What is the major disadvantage of the Two–phase commit protocol in distributed databases? How it is overcome in Three–phase commit protocol? **[8]**

OR

Q4) a) What are the different approaches to store a relation in the distributed database.Explain them in brief. [4]

b) Explain any two locking protocols with respect to distributed databases. [6]

c) Discuss the different system failures modes in distributed system. [8]

Q5) a) Explain in detail the XML document. [8]

b) Write short notes on : [8]
   i) SOAP.
   ii) Client – Server architecture.

OR

Q6) a) Explain in detail 3–Tier architecture. [8]

b) Write short notes on : [8]
   i) Domain specific DTD.
   ii) Querying XML data.

SECTION - II

Q7) a) What is a data cube? Explain any two operations on data cubes. [8]

b) Discuss the different ways of handling missing values in data cleaning.[8]

OR

Q8) a) What is meant by OLAP? Explain in brief. [6]

b) Discuss the different data smoothing techniques. [8]

c) What is meant by ETL Tool? [2]

Q9) a) What are Bayesian classifiers? [2]

b) State and explain K-means algorithm for clustering. [8]

c) What is the difference between descriptive and predictive data mining? [2]

d) Explain outlier analysis. [4]

OR

*Q10)*a) Explain the Market Basket analysis in brief. [5]

b) What is a decision tree? How are decision trees used for classification? Why are decision tree classifiers so popular? [8]

c) Explain Text mining in brief. [3]

Q11) a) Explain in the detail the measuring of the retrieval effectiveness. [8]

b) Explain the following terms : [10]

    i) Term frequency.

    ii) Relevance.

    iii) Proximity.

    iv) Concept–based querying.

    v) Stop words

OR

Q12) a) Explain in detail popularity ranking. [8]

b) Explain the following terms. [10]

    i) Web crawlers.

    ii) Page Rank.

    iii) Full text retrieval.

    iv) Inverse document frequency.

    v) Homonyms.

# # # #