Total No. of Questions : 12]

SEAT No. : ____

P2972

[Total No. of Pages : 4

[5354]-187
B.E. (Computer Engineering)
ADVANCE DATABASES
(2008 Pattern) (Elective - III)

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

Instructions to the candidates:

1) ***Answer three questions from Section - I and three questions from Section - II.***
2) ***Answers to the two sections should be written in separate books.***
3) ***Neat diagrams must be drawn wherever necessary.***
4) ***Figures to the right indicate full marks.***
5) ***Assume suitable data, if necessary.***

SECTION - I

Q1) a) Explain design issues of parallel system. **[6]**

b) Describe the benefits and drawbacks of pipelined parallelism. **[4]**

c) Describe a good way to parallelize each of the following: **[6]**

i) Selection
ii) Projection
iii) Aggregation

OR

Q2) a) Explain different partitioning techniques. **[6]**

b) Differentiate between interquery parallelism and intraquery parallelism **[6]**

c) Write a short note on parallel query optimization. **[4]**

Q3) a) Explain the difference between data replication in a distributed and the maintenance of a remote backup site. **[8]**

b) Explain Data fragmentation in distributed databases. **[6]**

c) Explain LDAP Data model. **[4]**

OR

P.T.O.

Q4) a) Explain in detail deadlock handing in distributed databases. **[8]**

b) Give example where the read one, write all available approach leads to an erroneous state. **[6]**

c) Explain any two Locking protocols with respect to distributed databases. **[4]**

Q5) a) Explain 3-tier client-server architecture in detail. **[8]**

b) Write short note on Web services. **[6]**

c) What is XSLT? **[2]**

OR

Q6) a) Explain the components of an XSL document with suitable example. **[8]**

b) Consider the following recursive DTD **[8]**

```
<! DOCTYPE parts [
<!ELEMENT parts (name, subpart info*)>
<!ELEMENT subpart info (part, quantity)>
<!ELEMENT name (# PCDATA)>
<!ELEMENT quality (# PCDATA)>
]>
```

Create a scheme in XSL scheme corresponding to this DTD.

SECTION - II

Q7) a) Explain multidimensional data models in details. **[8]**

b) Explain the following **[8]**

i) OLAP

ii) Data Cube

OR

Q8) a) Explain CUBE and ROLL-UP extended aggregation with suitable example. **[8]**

b) Explain the architecture of data warehouse and also explain different indexing technique used in data warehouse. **[8]**

Q9) a) What are different data cleaning methods? Explain binning & outlier analysis. **[6]**

b) State and explain K-MEANS algorithm for clustering. **[6]**

c) Consider the following data set. **[6]**

Food Item	Protein content	Fat Content
Fl	1.1	60
F2	8.2	20
F3	4.2	35
F4	1.5	21
F5	7.6	15
F6	2.0	55
F7	3.9	39

Find the cluster for the object in the dataset by using K-means algorithm ,if k = 4.

OR

Q10)a) What is Best split? Explain 1D3 algorithm to create decision tree. **[6]**

b) Explain the following: **[6]**

i) Text mining

ii) GINI index

iii) Information gain

c) Find the association rule for the given dataset which satisfy following requirements **[6]**

1) Support = 30%

2) Confidence = 90%

Customer	Products			
Cl	S1		S3	
C2		S2		
C3				S4
C4		S2	S3	S4
C5		S2	S3	
C6		S2	S3	
C7	S1	S2	S3	S4
C8	S1		S3	
C9	S1	S2	S3	
C10	S1	S2	S3	

Q11)a) What you mean by relevance ranking? Explain different methods of relevance ranking for the boolen & ranked query. **[8]**

b) Explain the following: **[8]**

i) Inverted Index

ii) Ontologies

iii) Stop words

OR

Q12)a) What is page ranking and popularity ranking? Explain in brief. **[6]**

b) Explain the following terms : **[10]**

i) Web crawlers.

ii) Homonyms.

iii) Vector space model.

iv) Zipfian distribution.