RV COLLEGE OF ENGINEERING

(An Autonomous Institution affiliated to VTU, Belagavi)

I Semester Master of Technology (Computer Science and Engineering)

ADVANCES IN DATABASE MANAGEMENT AND MINING

Time: 03 Hours *Maximum Marks: 100*

Instructions to candidates:

1. Each unit consists of two questions of 20 marks each.
2. Answer FIVE full questions selecting one from each unit (1 to 5).

UNIT-1

1	a	Design an Entity-Relationship diagram for the *MOVIE* database in which data is recorded about the movie industry. The data requirements are summarized as follows: i) Each movie is identified by title and year of release. Each movie has length in minutes. Each has a production company and each is classified under one or more genres(such as horror, action, comedy and so forth). Each movie has one or more directors and one or more actors appear in it. Each movie also has a plot outline. Finally, each movie has zero or more quotable quotes, each of which is spoken by a particular actor appearing in the movie. ii) Actors are identified by name and date of birth and appear in one or more movies. Each actor has a role in the movie. iii) Directors are also identified by name and date of birth and direct one or more movies. It is possible for a director to act in a movie (including one that he or she may also direct). iv) Production companies are identified by name and each has an address. A production company produces one or more movies.	10
	b	What is *XML*? What are its applications in database? What is stored in *XML*? Give an example.	10
		OR	
2	a	Why would you choose a database system instead of simply storing data in operating system files? When would it make sense not to use a database system?	10
	b	Considering any database of your choice explain the 7 steps in *ER* to Relational mapping.	10

UNIT-2

3	a	Briefly explain the object relational features of *SQL* with relevant *SQL* statements.	12
	b	Compare and contrast *ODL* and *ODMG*. Also write the significance of each one in detail.	08
		OR	
4	a	For extended type systems identify the implementation and related issues.	12
	b	What are the differences between *OQL* and *SQL*? Also write the features of *OQL*.	08

UNIT-3

5	a	With a neat diagram explain mediators. Considering a view in a relation and explain the mediators and wrappers.	10
	b	Describe how a typical lock manager is implemented. Why must lock and unlock be atomic operations? What is the difference between a lock and a latch? What are convoys and how should a lock manager handle them?	10
		OR	
6	a	Differentiate between the following with relevant diagram and examples. i) Federated and Data ware house ii) Parallel and Distributed database.	10
	b	Why data distribution needed? Explain with suitable examples?	10

UNIT-4

7	a	What are the differences between the three main types of data warehouse usage: information processing, analytical processing, and data mining? Discuss the motivation behind *OLAP* mining (*OLAM*).	10
	b	A database has five transactions. Let min sup = 60% and min conf = 80%. Final all frequent itemsets using Apriori algorithm. Table 1 *TID* | *items_bought* *T*100 | {*M*, *O*, *N*, *K*, *E*, *Y*} *T*200 | {*D*, *O*, *N*, *K*, *E*, *Y*} *T*300 | {*M*, *A*, *K*, *E* } *T*400 | {*M*, *U*, *C*, *K*, *Y*} *T*500 | {*C*, *O*, *O*, *K*, *I*, *E*}	10
		OR	
8	a	Clustering has been popularly recognized as an important data mining task with broad applications. Give one application example for each of the following cases: i) An application that takes clustering as a major data mining function. ii) An application that takes clustering as a preprocessing tool for data preparation for other data mining tasks.	10
	b	Using the given data in Question No.7 (b), find all Frequent Item set using *FP* growth algorithm.	10

UNIT-5

9	a	List out the challenges in mobile databases. With a general mobile computing architecture explain in detail.	10
	b	List and explain types of Multimedia data in current systems.	10
		OR	
10	a	Identify the scope of Geographical Information Systems. How can *GIS* be categorized? List the applications of each of them.	10
	b	Write a note on Genomics data management. Also list the existing biological databases.	10