Program: **MCA**

Curriculum Scheme: MCA (2year – 2020 Course)

Examination: M.C.A Semester I

Course Code: MCA13 and Course Name: Advanced Database Management System

Time: 2 hour 30 minutes   Max. Marks: 80

==

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	------- is a logical pointer to a row object
Option A:	ADT
Option B:	ref
Option C:	Both A and B
Option D:	None of these
2.	An itemset whose support is greater than or equal to a minimum support threshold is _____
Option A:	Itemset
Option B:	Frequent Itemset
Option C:	Infrequent Itemset
Option D:	Threshold Value
3.	A data warehouse is......................
Option A:	Updated by end users.
Option B:	Contains numerous naming conventions and formats
Option C:	Organized around important subject areas
Option D:	contain only current data
4.	Storing separate copy of the database at multiple location is _____-
Option A:	Data Replication
Option B:	Horizontal Fragmentation
Option C:	Vertical Fragmentation
Option D:	Horizontal and vertical Fragmentation
5.	Decision tree is
Option A:	Classification
Option B:	Prediction
Option C:	Both A and B
Option D:	None of these
6.	------ is an extension of the relational data model by including object orientation and constructs to deal with added data types.
Option A:	DBMS
Option B:	RDBMS
Option C:	ORDBMS
Option D:	OODBMS
7.	-------- is a program that traverses the hypertext structure in the Web.
Option A:	Harvest system

Option B:	Crawler
Option C:	Web log
Option D:	Search Engines
8.	What do you mean by support(A)?
Option A:	Total number of transactions containing A
Option B:	Total number of transactions not containing A
Option C:	Number of transactions containing A / Total number of transaction
Option D:	Number of transactions not containing A / Total number of transaction
9.	Agglomerative clustering follows ___
Option A:	Left to right
Option B:	Top to bottom
Option C:	Right to left
Option D:	Bottom to up
10.	The operation of moving from finer granular data to coarser granular data is called
Option A:	Roll up
Option B:	Roll down
Option C:	Reduction
Option D:	slice

	Solve any Two Questions out of Three. Each question carries 10 marks.
Q2.A	Define distributed database. Also explain the architecture of distributed database in detail.
Q2.B	Explain different pre-processing techniques in detail
Q2.C	Explain text mining and discuss in brief the Information retrieval methods.
	Solve any Two Questions out of Three. Each question carries 10 marks.
Q3.A	Define OLAP. Explain the different OLAP models with a suitable diagram.
Q3.B	What is Market Basket Analysis? Find out frequent itemsets and strong association rule from the given transaction using Apriori Algorithm with Min_Support of 50% and Confidence of 70%.

TID	Items
T1	Bread, Cheese, Juice, Eggs
T2	Bread, Cheese, Juice
T3	Bread, Milk, Yogurt
T4	Bread, Juice, Milk
T5	Cheese, Juice, Milk
T6	eggs, Milk, Yogurt
T7	Bread, Cheese, Juice
T8	eggs, Cheese, Juice
T9	Bread, Cheese, Juice
T10	Bread, yogurt, milk

Q3.C	Apply naïve baye's algorithm and predict the Class for unseen sample {chills=Yes, runny_nose=No,headache=mild,fever=Yes}

Chills	Runny_nose	Headache	Fever	Flu
Yes	No	mild	Yes	No
Yes	Yes	No	No	NO
Yes	No	strong	Yes	Yes
No	Yes	mild	Yes	Yes
No	Yes	mild	No	Yes
No	No	No	Yes	No
No	No	strong	No	No
No	No	strong	Yes	Yes
Yes	No	strong	yes	No
Yes	Yes	No	No	Yes

	Solve any Two Questions out of Three. Each question carries 10 marks.
Q4.A	Discuss K Nearest Neighbor algorithm with a suitable example
Q4.B	Explain the terms Entropy and Information Gain with an example.
Q4.C	Define Clustering. Consider seven objects with two attribute (A and B).Generate the Clusters using K-mean Clustering (K=3)

Object	A	B
A	1	1
B	1.5	2
C	3	4
D	5	7
E	3.5	5
F	4.5	5
G	3.5	4.5