UNIT 4			
Cluster Analysis Introduction-Requirements and overview of different categories-Partitioning method: Introduction - k-means-k-medoids-Hierarchical method: Introduction-Agglomerative vs. Divisive method-Distance measures in algorithmic methods-BIRCH technique-DBSCAN technique-STING technique-CLIQUE technique-Evaluation of clustering techniques			
PART-A (Multiple Choice Questions)			
Q. No	**Questions**	**Course Outcome**	**Competence BT Level**
1	_____ can be utilized to arrange the indexed lists into gatherings and present the outcomes in effectively available manner. A. Clustering B. Classification C. Association D. Pattern Mining **Answer: A**	CO4	BT 4
2	_____also helps in identification of common land use within a network for earth observation. This also helps in distinguishing housing classes in a community by type of home, size and geographic location. A. Clustering B. Classification C. Association D. Pattern Mining **Answer: A**	CO4	BT 4
3	Partitioning methods uses _____ that endeavors to improve the dividing by moving articles starting from one gathering to next. A. Recursive relocation B. Iterative relocation C. Progressive relocation D. Dynamic relocation **Answer: B**	CO4	BT 4
4	The time complexity of the *k*-means algorithm is_____. A. O(n) B. O(nkt) C. O(nk) D. O(nlogt) **Answer: B**	CO4	BT 5
5	An agglomerative various leveled grouping technique utilizes a _____strategy. A. bottom-up B. top-down C. iterative D. recursive **Answer: A**	CO4	BT 4

6	Divisive hierarchical clustering method uses a _____strategy. A. bottom-up B. top-down C. iterative D. recursive **Answer: B**	CO4	BT 4
7	_____is a tree structure that generally speaks to the procedure of various leveled clustering. A. Dendrogram B. CF Tree C. B+ Tree D. Balanced B+ **Answer: A**	CO4	BT 4
8	The grouping procedure is ended when the most extreme separation between closest bunches surpasses a client characterized edge, it is called a_____. A. Single linkage B. Complete linkage C. Double linkage D. Closed linkage **Answer: B**	CO4	BT 4
9	_____and _____parameters implicitly manage the ensuing cluster feature tree's size. A. Balancing Factor, Time B. Branching Factor , Time C. Branching Factor , Threshold D. Balancing Factor, Threshold **Answer: C**	CO4	BT 4
10	The time for BIRCH procedure _____. A. O(log n) B. O(n log n) C. O(n) D. $O(n^2)$ **Answer: C**	CO4	BT 5
11	The complexity of computation in DBSCAN _____ A. O(log n) B. O(n log n) C. O(n) D. $O(n^2)$ **Answer: B**	CO4	BT 5
12	The query processing time for STING _____. A. O(log g) B. O(g log g) C. O(g) D. $O(g^2)$ **Answer: C**	CO4	BT 5

13	In CLIQUE identifying the candidate search space is based on _____. A. Preprocessing data B. Apriori Property C. Frequent Pattern D. Partitioning **Answer: B**	CO4	BT 4
14	Choose the statistical method for measuring randomness of a variable. A. Hopkins Statistic B. Homogeneous hypothesis C. Non homogeneous hypothesis D. Probabilistic statistics **Answer: A**	CO4	BT 4
15	Clustering _____ assesses the plausibility of clustering breaks down on an informational collection and the nature of the grouping results. A. Validation B. Evaluation C. Verification D. Quality **Answer: B**	CO4	BT 4
16	Choose the wrong statement. A. K-means is same as k-nearest neighbor B. k-means aims at k clusters from n objects C. k-means uses vector quantization D. K-means is easy to partition the data set. **Answer: A**	CO4	BT 4
17	Which of the following clustering techniques requires a merging concept? A. Partitioning B. Hierarchical C. Naïve bayes D. Grid **Answer: B**	CO4	BT 4
18	What is Dendrogram? A. A tree illustrating the cluster arrangements in hierarchical technique B. A method of clustering C. A tree illustrating the cluster arrangements in partitioning technique D. A type of tree used in visualization technique **Answer: A**	CO4	BT 4
19	Which of the following tasks are correct on applications of clustering technique? a. Given a user's knowledge database, divide them into different market segments automatically. b. From the user's usage patterns on a website, identify different user groups. A. (a) only	CO4	BT 4

	B. both (a) and (b) C. (b) only D. Neither (a) nor (b) Answer: B		
20	Choose some of the considerations and requirements for cluster analysis. a. Required to incorporate user preference for cluster size and shape into clustering algorithm b. Needed Similarity measure between data objects a. Handling of various types of attributes like numeric, categorical b. Must know the number of output clusters apriori for all clustering algorithms A. a,b,c B. a,c,d C. b,c,d D. a,b,d **Answer: A**	CO4	BT 4
21	Choose the correct statement /statements. a. Clustering is unsupervised learning since it has no class labeled training data b. Must have all data objects ready before clustering c. Clustering has a wide range of applications in various fields like web search, multimedia, data summarization etc. d. In clustering we can put two dissimilar data objects into the same cluster. A. a, b B. a, c C. b, c D. a, b **Answer: C**	CO4	BT 4
22	The clustering algorithms most commonly evaluate clusters based on _____ A. Euclidean B. Manhattan C. ChebyChev's Distance D. Chi-Square **Answer: A**	CO4	BT 4
23	The K-means algorithm follows two steps iteratively in its loop. Choose the two steps. i. Assign points to its nearest cluster ii. Use of elbow method to choose value of k iii. Update the centroids of cluster based on current assessment iv. Validate the cluster formed. A. i , ii B. i ,iii C. ii,iii D. iv,iii	CO4	BT 4

	Answer: B		
24	Calculate the Euclidean Distance for D1 and D2 data points on a particular graph. D1= (2,0) D2=(1,3) A. 3.17 B. 2 C. 2.83 D. 4.48 **Answer: A**	CO4	BT 6
25	Given the cluster mean {1, 2, 3} is 2 and the mean {8, 9, 10, 25} is 13 for k=2. Compute the cluster variation. A. 189 B. 196 C. 209 D. 367 **Answer: B**	CO4	BT 6
26	Identify the inputs to the k-medoids algorithm. A. N: number of objects, Euclidean Distance B. K: no. of clusters, D:data set of n objects C. D: data set of n objects, Euclidean Distance D. K: no. of clusters, N: Number of objects **Answer: B**	CO4	BT 4
27	For Finding dissimilarities between two clusters in hierarchical clustering, which of the following metrics are used? 1. Single-Linkage 2. Complete-Linkage A. 1 Only B. 2 Only C. Both 1 and 2 D. Neither 1 nor 2 **Answer: C**	CO4	BT 4
28	_____is the resulting cluster shape in STING that can diminish the quality and precision of the clusters given the technique's quick processing time. A. Isothetic B. Spherical C. Oval D. Grid **Answer: A**	CO4	BT 4
29	Clustering is a type of learning by perception, as opposed to learning by models. Clustering is known as_____. A. Supervised Learning B. Unsupervised Learning C. Semi Supervised Learning D. Machine Learning	CO4	BT 4

	Answer: B		
30	Clustering decides groups dependent on Euclidean or Manhattan separation measures. Calculations dependent on such separation estimates will in general find _____. A. Spherical clusters B. Oval clusters C. Triangular clusters D. No shape in clusters **Answer: A**	CO4	BT 4