LAKIREDDY BALI REDDY COLLEGE OF ENGINEERING
(AUTONOMOUS)
Accredited by NAAC & NBA (CSE, IT, ECE, EEE & ME)
Approved by AICTE, New Delhi and Affiliated to JNTUK, Kakinada
L.B.Reddy Nagar, Mylavaram-521230, Krishna Dist, Andhra Pradesh, India
DEPARTMENT OF INFORMATION TECHNOLOGY

OBJECTIVE QUESTIONS

Name of Course Instructor: Michael Sadgun Rao Kona **Reg**: R20
Course Name & Code **:** Data Warehousing and Data Mining (**20CS10**) **Cycle**: 2
L-T-P Structure **: 3-0-0** **Credits:** 3
Program/Sem/Sec : B.Tech-IV Sem, Sections-A & B **A.Y.:**2021-22

Q. No	Question Description	Ans
1	Classification accuracy is A. A subdivision of a set of examples into a number of classes **B. Measure of the accuracy, of the classification of a concept that is given by a certain theory** C. The task of assigning a classification to a set of examples D. None of these	B
2	Some telecommunication company wants to segment their customers into distinct groups in order to send appropriate subscription offers, this is an example of A. Supervised learning B. Data extraction C. Serration **D. Unsupervised learning**	D
3	In the example of predicting number of babies based on storks' population size, number of babies is **A. outcome** B. feature C. attribute D. observation	A
4	Data independence means A. Data is defined separately and not included in programs B. Programs are not dependent on the physical attributes of data. C. Programs are not dependent on the logical attributes of data **D. Both (B) and (C).**	D
5	What is the relation between a candidate and frequent item sets? (a) A candidate itemset is always a frequent itemset **(b) A frequent itemset must be a candidate itemset** © No relation between these two (d) Strong relation with transactions	B
6	Which of the following is not a frequent pattern mining algorithm? (a) Apriori (b) FP growth © **Decision trees** (d) Eclat	C

7	What will happen if support is reduced? (a) Number of frequent item sets remains the same **(b) Some item sets will add to the current set of frequent item sets** © Some item sets will become infrequent while others will become frequent (d) Some item sets will become infrequent	B
8	Frequency of occurrence of an itemset is called as _____	Support Count
9	An itemset whose support is greater than or equal to a minimum support threshold is _____	Frequent Itemset
10	What does FP growth algorithm do? (a) It mines all frequent patterns through pruning rules with lesser support (b) It mines all frequent patterns through pruning rules with higher support **(c) It mines all frequent patterns by constructing a FP tree** (d) It mines all frequent patterns by constructing an itemsets	C
11	Which of the following is the direct application of frequent itemset mining? (a) Social Network Analysis **(b) Market Basket Analysis** (c) Outlier Detection (d) Intrusion Detection	B
12	What is association rule mining? (a) Same as frequent itemset mining **(b) Finding of strong association rules using frequent itemsets** (c) Using association to analyze correlation rules (d) Finding Itemsets for future trends	B
13	Hierarchical clustering should be mainly used for exploration. (A) True (B) False	True
14	Which is needed by K-means clustering? (A). defined distance metric (B). number of clusters (C). initial guess as to cluster centroids **(D). all of these**	D
15	Which function is used for k-means clustering? (**A). k-means** (B). k-mean (C). heatmap (D). none of the mentioned	A

16	Which clustering technique requires a merging approach? (A). Partitional **(B). Hierarchical** (C). Naive Bayes (D). None of the mentioned	B
17	._____consider the clusters as the dense region having some similarity and different from the lower dense region of the space **A) Density-Based** B) Hierarchical Based C) Grid-based D) None of these	A
18	What can we use in Hierarchical Clustering to find the right number of clusters ? A. The Elbow Method B. Decision Trees C. **Dendrograms** D. Histograms	C
19	In Python, what is the class used to fit hierarchical clustering to a data set ? A. HierarchicalClustering B. HClustering C. **AgglomerativeClustering** D. AgglomerativeHierarchical	C
20	If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction containjam, 10000 transaction contain both bread and jam. Then the support of bread and jam is _____. **A. 2%** B. 20% C. 3% D. 30%.	A
21	The _____ step eliminates the extensions of (k-1)-item sets which are not found to be frequent, from being considered for counting support.	Pruning
22	The _____ algorithm is based on the observation that the frequent sets are normally very fewin number compared to the set of all itemsets	Partition
23	_____is a metric to measure how often a randomly chosen element would be incorrectly identified. A) Information Gain **B) Gini Index** C) Entropy D) none of these	B
24	_____is the measure of uncertainty of a random variable, it characterizes the impurity of an arbitrary collection of examples. A) Information Gain B) Gini Index **C) Entropy** D) none of these	C
25	_____ refers to a model that can neither model the training data nor generalize to new data. A. good fitting	

	B. overfitting C. **underfitting** D. all of the above	C
26	Spam Classification is an example for? A. **Naive Bayes** B. Probabilistic condition C. Random Forest D. d. All the Above	A
27	Probabilistic Model of data within each class is? A. Discriminative classification B. Generative classification C. Probabilistic classification D. **Both b and c**	D
28	Which of the following is not clustering method? A) Density-Based B) Hierarchical Based C) Grid-based **D) Project Based**	D
29	Hierarchical Based Methods Consist of which category? A) Divisive B) Agglomerative **C) both a and b** D) None of these	C
30	_____clusters formed in this method forms a tree-type structure based on the hierarchy. A) Density-Based **B) Hierarchical Based** C) Grid-based D) None of these	B

Title	Course Instructor	Module Coordinator	Head of the Department
Name of the Faculty	**Michael Sadgun Rao Kona**		**Dr.B.Srinivasa Rao**
Signature			