Program: BE **Information Technology** Engineering

Curriculum Scheme: Revised **2012**

Examination: **Third Year** Semester **VI**

Course Code: **TEITC604** and Course Name: **DMBI**

Time: 1 hour Max. Marks: 50

==============================================================================

Note to the students:- All the Questions are compulsory and carry equal marks .

|  |  |
| --- | --- |
| Q1.  | Which of the following is Data Mining Tool  |
| Option A: | Weka |
| Option B: | Borland C |
| Option C: | C/C++ |
| Option D:  | Visual C |
|  |  |
| Q2. | Nominal Attributes are also referred as |
| Option A: | Ordinal Attributes |
| Option B: | Interval -Scaled Attributes |
| Option C: | Ratio Scaled Attributes |
| Option D: | Categorical Attributes |
|  |  |
| Q3. | Which among the following attributes calculates ranking or meaningful order of data |
| Option A: | Categorical Attributes |
| Option B: | Numeric Attributes |
| Option C: | Ratio-Scaled Attributes |
| Option D: | Ordinal Attributes |
|  |  |
| Q4. | The mode for set of data is ------------ |
| Option A: | Lowest value of data set |
| Option B: | Mid value of data set |
| Option C: | A value that occurs most frequently in a data set |
| Option D: | Highest value of data set |
|  |  |
| Q5. | In Knowledge Discovery process of Data Mining, Data Integration------- |
| Option A: | Removes noise and inconsistent data |
| Option B: | Combines multiple data sources |
| Option C: | Performs analysis on data |
| Option D:  | Identifies truly interesting patterns from the data |
|  |  |
| Q6. | Use of Measure of Central Tendency for the attribute to fill in the missing value is part of which process |
| Option A: | Data Cleaning |
| Option B: | Data Integration |
| Option C: | Data Reduction |
| Option D:  | Data Transformation |
|  |  |
| Q7.  | Taxonomy Formation term is associated with |
| Option A: | Classification |
| Option B: | Outlier Analysis |
| Option C: | Cluster Analysis |
| Option D:  | Association Analysis |
|  |  |
| Q8.  | Regression analysis in classification is |
| Option A: | A methodology mostly used for numeric prediction |
| Option B: | A methodology mostly used for binary data prediction |
| Option C: | A methodology mostly used for ordinary data prediction |
| Option D:  | A methodology mostly used for ordinary and binary data prediction |
|  |  |
| Q9. | The Bayes rule can be used to |
| Option A: | Solve Queries |
| Option B: | Answering Probabilistic Query |
| Option C: | Increasing Complexity |
| Option D:  | Decreasing Complexity |
|  |  |
| Q10.  | Sensitivity is measure of |
| Option A: | True Positive Recognition rate which is Proportion of positive tuples that are correctly identified |
| Option B: | The Negative Recognition rate which is the proportion of negative tuples that are correctly identified |
| Option C: | Percentage of test set tuples that are correctly classified |
| Option D:  | Percentage of errors made over the whole set of instances used for testing |
|  |  |
| Q11.  | K-fold Cross Validation is  |
| Option A: | Linear in K |
| Option B: | Quadratic in K |
| Option C: | Cubic in K |
| Option D:  | Exponential in K |
|  |  |
| Q12.  | What is association rule mining? |
| Option A: | Same as frequent itemset mining |
| Option B: | Finding of strong association rules using frequent itemsets |
| Option C: | Using association to analyse correlation rules |
| Option D: | same as FP growth algorithm |
|  |  |
| Q13. | k-means clustering is also referred to as \_\_\_\_\_? |
| Option A: | Non-hierarchical clustering |
| Option B: | Optimizing partitioning |
| Option C: | Divisive clustering |
| Option D:  | Agglomerative clustering |
|  |  |
| Q14.  | Process of manage information about customers to maximize loyalty is said to be  |
| Option A: | Company relationship management |
| Option B: | Supplier management |
| Option C: | Retailer’s management |
| Option D:  | Customer relationship management |
|  |  |
| Q15. | In clustering \_\_\_\_? |
| Option A: | Groups are not predefined |
| Option B: | Groups are predefined |
| Option C: | Depends on the data |
| Option D:  | Supervised data mining |
|  |  |
| Q16.  | Which of the following algorithm is most sensitive to outliers? |
| Option A: | K-means clustering algorithm |
| Option B: | K-medians clustering algorithm |
| Option C: | K-modes clustering algorithm |
| Option D:  | K-medoids clustering algorithm |
|  |  |
| Q17. | What is clustering? |
| Option A: | Process of grouping similar objects |
| Option B: | Process of classifying new object |
| Option C: | Process of new object |
| Option D: | Process of deleting new object |
|  |  |
| Q18. | What are closed itemsets? |
| Option A: | An itemset for which at least one proper super-itemset has same support |
| Option B: | An itemset whose no proper super-itemset has same support |
| Option C: | An itemset for which at least super-itemset has same confidence |
| Option D:  | An itemset whose no proper super-itemset has same confidence |
|  |  |
| Q19.  | Why is correlation analysis important? |
| Option A: | To make apriori memory efficient |
| Option B: | To weed out uninteresting frequent itemsets |
| Option C: | To find large number of interesting itemsets |
| Option D:  | To restrict the number of database iterations |
|  |  |
| Q20. | Decision support system involves all of the following types of analytical modeling activities except? |
| Option A: | what-if analysis  |
| Option B: | Sensitivity analysis  |
| Option C: | Goal-seeking analysis |
| Option D: | Heuristics |
|  |  |
| Q21. | \_\_\_\_\_\_\_\_ is a performance management tool that recapitulates an organization’s performance from several standpoints on a single page.  |
| Option A: | Balanced Scorecard |
| Option B: | Data Cube |
| Option C: | Dashboard |
| Option D:  | Dice |
|  |  |
| Q22.  | The right-hand side of an association rule is called \_\_\_\_\_. |
| Option A: | onset. |
| Option B: | Consequent |
| Option C: | Antecedent |
| Option D:  | Precedent |
|  |  |
| Q23. | What should be a major characteristic of a DSS? |
| Option A: | Automates decision making |
| Option B: | Includes a spreadsheet model |
| Option C: | Responds quickly to the changing needs of decision makers |
| Option D:  | Manual decision making |
|  |  |
| Q24.  | Which clustering procedure is characterized by the formation of a tree like structure? |
| Option A: | Hierarchical clustering |
| Option B: | Optimizing partitioning |
| Option C: | Partition based clustering |
| Option D:  | Density clustering |
|  |  |
| Q25. | Advantage of CRM  |
| Option A: | Cost of the Software |
| Option B: | Improve overall relationship with customer |
| Option C: | Customization of the Business |
| Option D:  | Stalking, not wooing, customers |