University of Mumbai
Examination 2020 under cluster 4 (PCE)

Program: BE Computer Engineering
Curriculum Scheme: Rev2012
Examination: Final Year Semester VIII
Course Code: CPC801 and Course Name: Data Warehouse and Mining

Time: 1hour   Max. Marks: 50

===

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

Q1.	The core of the multidimensional model is the .... , which consists of a large set of facts and a number of dimensions
Option A:	Data Cube
Option B:	Data Model
Option C:	Multidimensional cube
Option D:	Dimensions cube
Q2.	The data from the operational environment are .... of data warehouse
Option A:	Older detail data
Option B:	Lightly Summarized data
Option C:	Current detail data
Option D:	Highly summarized data
Q3.	The .... exposes the information being captured, stored, and managed by operational systems.
Option A:	Data warehouse view
Option B:	Data source view
Option C:	Top-down view
Option D:	Business query view
Q4.	Which of the following is not a kind of data warehouse application?
Option A:	Transaction Processing
Option B:	Data Mining
Option C:	Analytical processing
Option D:	Information processing
Q5.	Which one manages both current and historic transactions?
Option A:	OLTP
Option B:	OLAP
Option C:	Spread sheet
Option D:	XML
Q6.	An OLAP tool provides for ......
Option A:	Multidimensional analysis
Option B:	Rolling
Option C:	Slicing and dicing
Option D:	Rotation
Q7.	Which of the following tools a business intelligence system will have?

Option A:	ETL
Option B:	Data mining tool
Option C:	Reporting tool
Option D:	OLAP tool
Q8.	Which of the following is not an ETL tool?
Option A:	Informatics
Option B:	Oracle warehouse builder
Option C:	Datastage
Option D:	Visual studio
Q9.	What ETL Stands for ....
Option A:	Expansion, Transform, Load
Option B:	Transformation
Option C:	Extract, Transform, Load
Option D:	Data Flow
Q10.	What do data warehouses support?
Option A:	OLAP
Option B:	OLTP
Option C:	OLAP and OLTP
Option D:	Operational databases
Q11.	What Knowledge discovery in database
Option A:	whole process of extraction of knowledge from data
Option B:	selection of data.
Option C:	coding
Option D:	cleaning the data
Q12.	Which of the following activities is NOT a data mining task
Option A:	Predicting the future stock price of a company using historical records
Option B:	Monitoring and predicting failures in a hydropower plant
Option C:	Extracting the frequencies of a sound wave
Option D:	Monitoring the heart rate of a patient for abnormalities
Q13.	Data Visualization in mining cannot be done using
Option A:	Photos
Option B:	Graphs
Option C:	Charts
Option D:	Information Graphics
Q14.	Which is a comparison of the general features of the target class data objects against the general features of objects from one or multiple contrasting classes
Option A:	Data Characterization
Option B:	Data Classification
Option C:	Data discrimination

Option D:	Data selection
Q15.	Which of the following is not a data pre-processing method
Option A:	Data Visualization
Option B:	Data Discretization
Option C:	Data Cleaning
Option D:	Data Reduction
Q16.	Dimensionality reduction reduces the data set size by removing
Option A:	Composite attributes
Option B:	Derived attributes
Option C:	Relevant attributes
Option D:	Irrelevant attributes
Q17.	To detect fraudulent usage of credit cards, the following data mining task should be used
Option A:	Outlier analysis
Option B:	Prediction
Option C:	Association analysis
Option D:	Feature selection
Q18.	A _____ or tree graph is a graphical device for displaying clustering results. Vertical lines represent clusters that are joined together. The position of the line on the scale indicates the distances at which clusters were joined.
Option A:	Dendogram
Option B:	Scattergram
Option C:	Scree plot
Option D:	Icicle diagram
Q19.	Bayesian classifiers is
Option A:	A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.
Option B:	Any mechanism employed by a learning system to constrain the search space of a hypothesis
Option C:	An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation.
Option D:	Computational procedure that takes some value as input and produces some value as output
Q20.	Classification is
Option A:	A subdivision of a set of examples into a number of classes
Option B:	A measure of the accuracy, of the classification of a concept that is given by a certain theory
Option C:	The task of assigning a classification to a set of examples
Option D:	Systems that can be used without knowledge of internal operations

Q21	Which of the following is direct application of frequent itemset mining?
Option A:	Social Network Analysis
Option B:	Market Basket Analysis
Option C:	Outlier Detection
Option D:	Intrusion Detection
Q22	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 transactions
Option D:	Number of transactions not containing A / Total number of transactions
Q23	When do you consider an association rule interesting?
Option A:	If it only satisfies min_support
Option B:	If it only satisfies min_confidence
Option C:	If it satisfies both min_support and min_confidence
Option D:	There are other measures to check so
Q24	MOLAP stands for......
Option A:	Multidimensional Online Analytical Product
Option B:	Multiorder Online Analytical Product
Option C:	Multidimensional Online Analytical Processing
Option D:	Multifunctional Online Analytical Processing
Q25	What is of the following is not information delivery component?
Option A:	Ad hoc Report
Option B:	Data Mining
Option C:	Production Data
Option D:	Statistical Analysis

University of Mumbai

Examination 2020 under cluster 4 (PCE)

Program: BE Computer Engineering
Curriculum Scheme: Rev2012
Examination: Final Year Semester VIII
Course Code: CPC801 and Course Name: Data Warehouse and Mining

Time: 1hour   Max. Marks: 50

==

Question	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	A
Q2.	C
Q3.	B
Q4	A
Q5.	B
Q6.	C
Q7.	D
Q8.	D
Q9.	C
Q10.	A
Q11.	A
Q12.	C
Q13.	A
Q14.	C
Q15.	A
Q16.	D
Q17.	A
Q18.	A
Q19.	A
Q20.	A
Q21	B
Q22	C
Q23	C
Q24	C
Q25	C