

University of Mumbai
Examination 2020 under cluster 4 (PCE)

Program: BE Computer Engineering
Curriculum Scheme: Rev2016
Examination: Third Year Semester VI
Course Code: CSC603 and Course Name: DWM

Time: 1 hour

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.	The following is not the data extraction issue
Option A:	Source identification
Option B:	Time window
Option C:	Report generation
Option D:	Job sequencing
Q6.	The data capture through database triggers is the ---- data extraction.
Option A:	Deferred
Option B:	Immediate
Option C:	Replication technology
Option D:	future
Q7.	The following step is not part of the KDD process

University of Mumbai
Examination 2020 under cluster 4 (PCE)

Option A:	Data Mining
Option B:	Preprocessing
Option C:	Trasformation
Option D:	Extraction
Q8.	The following is not the major steps in ETL process
Option A:	Write procedure for all data load
Option B:	Plan for aggregate table
Option C:	Generate report
Option D:	Establish comprehensive data extraction rule
Q9.	The operation of moving from finer granular data to coarser granular data is called
Option A:	Roll up
Option B:	Drill down
Option C:	Pivot
Option D:	Slicing
Q10.	Which of the following is not a data mining task?
Option A:	Feature Subset Detection
Option B:	Association Rule
Option C:	Regression
Option D:	Sequential pattern discovery
Q11.	Which of the following is an Entity identification problem?
Option A:	One person with different email address
Option B:	One person's name written in different way
Option C:	Title for person
Option D:	person with multiple phone numbers
Q12.	In Binning, we first sort data and partition into (equal-frequency) bins and then which of the following is not a valid step
Option A:	smooth by bin boundaries
Option B:	smooth by bin median
Option C:	smooth by bin means
Option D:	smooth by bin values
Q13.	Incorrect or invalid data is known as
Option A:	Missing data
Option B:	Changing data
Option C:	Outlier
Option D:	Noisy data
Q14.	Which of the following is NOT a data quality related issue?
Option A:	Missing values
Option B:	Outlier records

University of Mumbai
Examination 2020 under cluster 4 (PCE)

Option C:	Duplicate records
Option D:	Attribute value range
Q15.	Which stage of data mining involves preparation and collection of data?
Option A:	Validation
Option B:	Collection
Option C:	Exploration
Option D:	Selection
Q16.	Data Mining is the step between--- and ----.
Option A:	Data Transformation, Data Selection
Option B:	Data Transformation, Pattern Evaluation
Option C:	Pattern Evaluation, Knowledge Presentation
Option D:	Data Cleaning, Data Integration
Q17.	The following is not the major Data Mining issues
Option A:	Diverse Data Types Issues
Option B:	Performance Issues
Option C:	Dimension Modeling Issue
Option D:	Mining Methodology and User Interaction
Q18.	Noise is
Option A:	A component of a network
Option B:	In the context of KDD and data mining, this refers to random errors in a database table.
Option C:	One of the defining aspects of a data warehouse
Option D:	A measure of the desired maximal complexity of data mining algorithms
Q19.	A _____ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.
Option A:	Decision tree
Option B:	Graphs
Option C:	Trees
Option D:	Neural Networks
Q20.	Which of the following is finally produced by Hierarchical Clustering?
Option A:	final estimate of cluster centroids
Option B:	tree showing how close things are to each other
Option C:	assignment of each point to clusters
Option D:	defined distance metric
Q21.	The ID3 algorithm is used to build a
Option A:	decision tree
Option B:	map

University of Mumbai
Examination 2020 under cluster 4 (PCE)

Option C:	clusters
Option D:	itemsets
Q22.	Which is not a classification technique ?
Option A:	Regression
Option B:	Decision trees
Option C:	Neural Networks
Option D:	k-means
Q23.	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
Q24.	Which of the following is related to GIS
Option A:	Euclidean space
Option B:	Ramanujan space
Option C:	Pythagorian space
Option D:	Manhattan space
Q25.	Postpruning is _____
Option A:	prune after building the tree
Option B:	prune while building the tree
Option C:	remove exactly one branch
Option D:	remove exactly two branches

University of Mumbai
Examination 2020 under cluster 4 (PCE)

Program: BE Computer Engineering

Curriculum Scheme: Rev2016

Examination: Third Year Semester VI

Course Code: CSC603 and Course Name: DWM

Time: 1 hour

Max. Marks: 50

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