

MCQ_TE-B_Subject-DWM_R16_Block2

The name, username and photo associated with your Google account will be recorded when you upload files and submit this form.

Not ameythakur@ternaengg.ac.in? [Switch account](#)

* Required

MCQ

Q1 *

1.	The purpose of the operational system is used to _____
Option A:	Run the business in real time and is based on historical data
Option B:	Takes strategic decisions for business
Option C:	Support decision making and is based on historical data
Option D:	Run the business in real time and is based on current data

☐ A

☐ B

☐ C

☒ D

Q2 *

2.	Which of following describes a data warehouse well?
Option A:	Can be updated by end users.
Option B:	Contains numerous naming conventions and formats.
Option C:	Organized around important subject areas.
Option D:	Contains only current data

☐ A



☐ B

☒ C

☐ D

Q3 *

3.	Expected amount of information (in bits) needed to assign a class to a randomly drawn object is _____
Option A:	Gain ratio
Option B:	<u>Gini Index</u>
Option C:	Entropy
Option D:	Information Gain

☐ A

☐ B

☒ C

☐ D

Q4 *

4.	Which of the following achieves data reduction by detecting redundant attributes
Option A:	Data cube aggregation
Option B:	Dimension reduction
Option C:	Data compression
Option D:	<u>Numerosity reduction</u>

☐ A

☒ B

☐ C

☐ D



Q5 *

5.	The fraudulent usage of credit card-can be detected using data mining task should be used
Option A:	Prediction
Option B:	Outlier analysis
Option C:	Association analysis
Option D:	Correlation

☐ A

☒ B

☐ C

☐ D

Q6 *

6. Given the record of users and movies viewed. Using Jaccard similarity measures, find similarity between {A-B,A-C,B-C }

Users	Movie 1	Movie 2	Movie 3	Movie 4	movie 5
A	1	0	1	0	1
B	0	0	1	0	1
C	0	1	0	0	1

Option A: {0.67,0.25,0.33}

Option B: {0.67,0.33,0.25}

Option C: {0.5,0.33,0.67}

Option D: {0.5,0.25,0.67}

☒ A

☐ B

☐ C



☐ D

Q7 *

7.	Five-number summary of a distribution (Minimum, Q1, Median, Q3, Maximum) is displayed by-----
Option A:	Histogram
Option B:	<u>quantile plot</u>
Option C:	Scatterplot
Option D:	Box plot

☐ A

☐ B

☐ C

☒ D

Q8 *

8.	If a set is a frequent set and no superset of this set is a frequent set, then it is called _____.
Option A:	maximal frequent set
Option B:	border set
Option C:	lattice
Option D:	infrequent sets

☒ A

☐ B

☐ C

☐ D

Q9 *

9.	_____ is a mining task that examines the web and hyperlinks structure that connect web pages.
Option A:	Web content mining
Option B:	Web structure mining
Option C:	Web usage mining
Option D:	Web link mining

☐ A

☒ B

☐ C

☐ D

Q10 *

10.	What does Web content mining involve?
Option A:	analyzing the universal resource locator in Web pages
Option B:	analyzing the unstructured content of Web pages
Option C:	analyzing the pattern of visits to a Web site
Option D:	analyzing the PageRank and other metadata of a Web page

☐ A

☒ B

☐ C

☐ D

Q11 *

11.	A sub-database which consists of set of prefix paths in the FP-tree co-occurring with the suffix pattern is called as
Option A:	Suffix path
Option B:	FP-tree
Option C:	Prefix path
Option D:	Condition pattern base

- ☐ A
- ☐ B
- ☐ C
- ☒ D

Q12 *

12.	In star schema, there is one fact table as F1 is connected with four-dimension tables D1, D2, D3, D4 then fact table will have how many foreign keys?
Option A:	2
Option B:	4
Option C:	3
Option D:	5

- ☐ A
- ☒ B
- ☐ C
- ☐ D



Q13 *

13.	If Mean salary is 54,000 Rs. and standard deviation is 16,000 Rs. then find z score value of 73,600 Rs. salary
Option A:	1.225
Option B:	0.351
Option C:	1.671
Option D:	1.862

☒ A

☐ B

☐ C

☐ D

Q14 *

14.	The generalization of cross-tab which is represented visually is _____ which is also called as <u>data cube</u> .
Option A:	Two-dimensional cube
Option B:	Multidimensional cube
Option C:	N-dimensional cube
Option D:	Cuboid

☒ A

☐ B

☐ C

☐ D



Q15 *

15.	In KDD and Data mining, noise is referred to as
Option A:	Complex data
Option B:	Meta data
Option C:	Error
Option D:	Repeated data

- ☐ A
- ☐ B
- ☒ C
- ☐ D

Q16 *

16.	Find the IQR of the data set {3, 7, 8, 5, 12, 14, 21, 13, 18}.
Option A:	6
Option B:	12
Option C:	16
Option D:	10

- ☐ A
- ☐ B
- ☐ C
- ☒ D



Q17 *

17.	Which of the following is not a method to estimate a classifier's accuracy
Option A:	Holdout method
Option B:	Random Sampling
Option C:	Information Gain
Option D:	Bootstrap

- ☐ A
- ☐ B
- ☒ C
- ☐ D

Q18 *

18.	For questions given below consider the data Transactions : T1 {F, A, D, B} T2 {D, A, C, E, B} T3 {C, A, B, E} T4 {B, A, D} With minimum support is 60% and the minimum confidence is 80%. Which of the following is not valid association rule?
Option A:	A -> B
Option B:	B -> A
Option C:	D -> A
Option D:	A -> D

- ☐ A
- ☐ B
- ☐ C
- ☒ D

Q19 *

19.	To calculate distance between two <u>isotheticrectangles</u> , _____ is efficient approach and produces cluster of high quality
Option A:	CLARA
Option B:	PAM
Option C:	Spatial mining
Option D:	IR Approximation

- ☐ A
- ☐ B
- ☐ C
- ☒ D

Q20 *

20.	Geographers typically model the world with objects located at different places on surface of the earth. Through _____ model, the real word entities are represented by lines, points and polygons
Option A:	Vector data model
Option B:	Raster data model
Option C:	Network data model
Option D:	Topology data model

- ☒ A
- ☐ B
- ☐ C
- ☐ D

Back

Submit



Never submit passwords through Google Forms.

Google Forms

