

USN : 

Adichunchanagiri Institute of Technology

AIT,Chikkamagaluru

Information Science And Engineering

I - Other Assessment (Seminar/quiz)

Semester: 6-CBCS 2021

Date: 10 Jul 2024

Subject: DATA MINING AND DATA WAREHOUSING (21IS643)

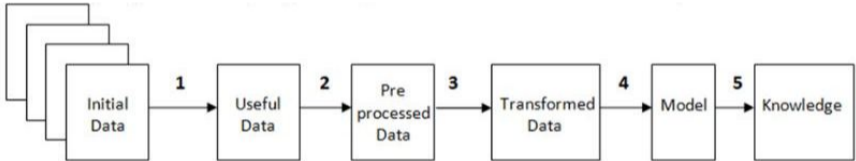
Faculty: Mr Abhishek N

Max Marks: 20

Answer All Questions

Q.No		Marks	CO	BT/CL
1	What is data mining? Select the correct answer 1). Deleting unnecessary data 2). Sorting data alphabetically 3). Storing data securely 4). Extracting useful patterns or information from large datasets	1	CO1	L1
2	Which of the following is not a basic data mining task? Select the correct answer 1). Spooling 2). Prediction 3). Classification 4). Clustering	1	CO2	L2
3	Which of the following is not an issue in data mining? Select the correct answer 1). High dimensionality 2). Shortage of data 3). Overfitting 4). Outliers	1	CO2	L2

4	<p>Which among the following are not among Various Operations in OLAP Cube? Select the correct answer</p> <ol style="list-style-type: none"> 1). Sticking 2). Dice 3). Drill down 4). Roll up 	1	CO1	L2
5	<p>Which of the following is true about bottom up discretization? Select the correct answer</p> <ol style="list-style-type: none"> 1). Some the values are treated as potential split points 2). All the values are treated as potential split points 3). Split points are not considered 4). Only one value are treated as potential split points 	1	CO2	L2
6	<p>The analysis of the data points that deviate from the general expected behavior of data in the data set is called _____ Select the correct answer</p> <ol style="list-style-type: none"> 1). Cluster analysis 2). Relevance analysis 3). Anomaly analysis 4). Regression analysis 	1	CO3	L1
7	<p>Data cleaning involves filling in missing values Select the correct answer</p> <ol style="list-style-type: none"> 1). True 2). False 	1	CO2	L1
8	<p>Data integration is not a step in data preprocessing Select the correct answer</p> <ol style="list-style-type: none"> 1). True 2). False 	1	CO2	L2

9	<p>Which of the following is not a form of data transformation? Select the correct answer</p> <ol style="list-style-type: none"> 1). Normalization 2). Discretization 3). Concept hierarchy 4). Compression 	1	CO2	L2
10	<p>In the diagram given below, depicting the KDD process, what is the correct order of tasks performed? Select the correct answer</p> <ol style="list-style-type: none"> 1). Selection, Preprocessing, Transformation, Data mining, Interpretation 2). Selection, Preprocessing, Transformation, Interpretation, Data mining 3). Interpretation, Preprocessing, Transformation, Data mining, Selection 4). Data mining, Preprocessing, Transformation, Selection, Interpretation 	1	CO2	L1
11	<p>Sometimes the data under analysis has too many attributes, some of which may not be useful as per the task being performed. This problem is referred to as _____ Select the correct answer</p> <ol style="list-style-type: none"> 1). Overfitting 2). Dimensionality curse 3). Outliers in data 4). Missing data 	1	CO3	L2
12	<p>What is the full form of KDD in the data mining process? Select the correct answer</p> <ol style="list-style-type: none"> 1). Knowledge data house 2). Knowledge data definition 3). Knowledge discovery data 4). Knowledge discovery database 	1	CO2	L2

13	<p>Among which of the following can be used by the warehouse? Select the correct answer</p> <ol style="list-style-type: none"> 1). Database table 2). Online database 3). Flat files 4). All of the above 	1	CO1	L2
14	<p>Total fact table the star schema is composed of? Select the correct answer</p> <ol style="list-style-type: none"> 1). 4 2). 3 3). 2 4). 1 	1	CO1	L2
15	<p>_____ is the heart of the warehouse Select the correct answer</p> <ol style="list-style-type: none"> 1). data mining database servers. 2). data warehouse database servers 3). data mart database servers 4). relational data base servers 	1	CO1	L2
16	<p>Which clustering algorithm does not require specifying the number of clusters in advance? Select the correct answer</p> <ol style="list-style-type: none"> 1). K-Means 2). DBSCAN 3). Agglomerative 4). Mean-Shift 	1	CO4	L2
17	<p>_____ is a good alternative to the star schema. Select the correct answer</p> <ol style="list-style-type: none"> 1). star schema. 2). snowflake schema 3). fact constellation 4). star-snowflake schema 	1	CO1	L2

18	<p>Fact tables are _____.</p> <p>Select the correct answer</p> <ol style="list-style-type: none">1). completely demoralized.2). partially demoralized.3). completely normalized.4). partially normalized.	1	CO1	L2
19	<p>Which of the following is a goal of clustering algorithms?</p> <p>Select the correct answer</p> <ol style="list-style-type: none">1). Classification2). Regression3). Dimensionality reduction4). Grouping similar data points together	1	CO4	L2
20	<p>Which clustering algorithm is based on the concept of centroids?</p> <p>Select the correct answer</p> <ol style="list-style-type: none">1). K-Means2). DBSCAN3). Agglomerative4). Mean-Shift	1	CO4	L2

DATA MINING AND DATA WAREHOUSING

9 Jul 2024

Totally 20 Marks

Total: 20 Points

🕒 20 minutes

* Required

1

USN *

2

NAME *

3

What is data mining * (1 Point)

- Deleting unnecessary data
- Sorting data alphabetically
- Storing data securely
- Extracting useful patterns or information from large datasets

4

Which of the following is not a basic data mining task? * (1 Point)

- Spooling
- Prediction
- Classification
- Clustering

5

Which of the following is not an issue in data mining? * (1 Point)

- High dimensionality
- Shortage of data
- Overfitting
- Outliers

6

Which among the following are not among Various Operations in OLAP Cube? * (1 Point)

- Sticking
- Dice
- Drill down
- Roll up

7

Which of the following is true about bottom up discretization? *

(1 Point)

- Some the values are treated as potential split points
- All the values are treated as potential split points
- Split points are not considered
- Only one value are treated as potential split points

8

The analysis of the data points that deviate from the general expected behavior of data in the data set is called _____ * (1 Point)

- Cluster analysis
- Relevance analysis
- Anomaly analysis
- Regression analysis

9

Data cleaning involves filling in missing values. * (1 Point)

- True
- False

10

Data integration is not a step in data preprocessing. * (1 Point)

- True
- False

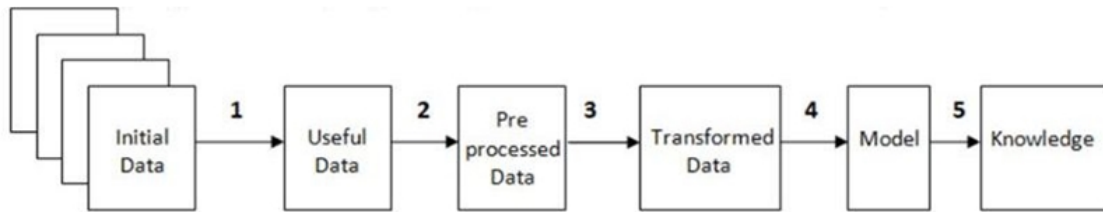
11

Which of the following is not a form of data transformation? *
(1 Point)

- Normalization
- Discretization
- Concept hierarchy
- Compression

12

In the diagram given below, depicting the KDD process, what is the correct order of tasks performed? * (1 Point)



- Selection, Preprocessing, Transformation, Data mining, Interpretation
- Selection, Preprocessing, Transformation, Interpretation, Data mining
- Interpretation, Preprocessing, Transformation, Data mining, Selection
- Data mining, Preprocessing, Transformation, Selection, Interpretation

13

Sometimes the data under analysis has too many attributes, some of which may not be useful as per the task being performed. This problem is referred to as _____ * (1 Point)

- Overfitting
- Dimensionality curse
- Outliers in data
- Missing data

14

What is the full form of KDD in the data mining process? * (1 Point)

- Knowledge data house
- Knowledge data definition
- Knowledge discovery data
- Knowledge discovery database

15

Among which of the following can be used by the warehouse? *
(1 Point)

- Database table
- Online database
- Flat files
- All of the above

16

Total fact table the star schema is composed of? * (1 Point)

- 4
- 3
- 2
- 1

17

 is the heart of the warehouse
* (1 Point)

- data mining database servers.
- data warehouse database servers
- data mart database servers
- relational data base servers

18

Which clustering algorithm does not require specifying the number of clusters in advance? * (1 Point)

- K-Means
- DBSCAN
- Agglomerative
- Mean-Shift

19

_____ is a good alternative to the star schema. * (1 Point)

- star schema.
- snowflake schema
- fact constellation
- star-snowflake schema

20

Fact tables are _____. * (1 Point)

- completely demoralized.
- partially demoralized.
- completely normalized.
- partially normalized.

21

Which of the following is a goal of clustering algorithms? * (1 Point)

- Classification
- Regression
- Dimensionality reduction
- Grouping similar data points together

Which clustering algorithm is based on the concept of centroids? *
(1 Point)

- K-Means
- DBSCAN
- Agglomerative
- Mean-Shift

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.

DATA MINING AND DATA WAREHOUSING

68 Responses 17.5 Average Score Active Status

1. NAME

68
Responses

Latest Responses

"Hamsa CS"

"Ananya HR"

"Gourav Shetti"

2. USN

68
Responses

Latest Responses

"4AI21IS020"

"4AI21IS005"

"4AI21IS017"

3. What is data mining (1 point)

97% of respondents (66 of 68) answered this question correctly.

- Deleting unnecessary data 2
- Sorting data alphabetically 0
- Storing data securely 0
- Extracting useful patterns or i... 66 ✓



4. **Which of the following is not a basic data mining task?** (1 point)

96% of respondents (65 of 68) answered this question correctly.

<input type="radio"/> Spooling	65	✓
<input type="radio"/> Prediction	2	
<input type="radio"/> Classification	0	
<input type="radio"/> Clustering	1	



5. **Which of the following is not an issue in data mining?** (1 point)

87% of respondents (59 of 68) answered this question correctly.

<input type="radio"/> High dimensionality	6	
<input type="radio"/> Shortage of data	59	✓
<input type="radio"/> Overfitting	1	
<input type="radio"/> Outliers	2	



6. **Which among the following are not among Various Operations in OLAP Cube?** (1 point)

93% of respondents (63 of 68) answered this question correctly.

<input type="radio"/> Sticking	63	✓
<input type="radio"/> Dice	3	
<input type="radio"/> Drill down	2	
<input type="radio"/> Roll up	0	



7. **Which of the following is true about bottom up discretization?** (1 point)

90% of respondents (61 of 68) answered this question correctly.

- Some the values are treated a... 5
- All the values are treated as p... 61 ✓
- Split points are not considered 0
- Only one value are treated as ... 2



8. **The analysis of the data points that deviate from the general expected behavior of data in the data set is called _____** (1 point)

96% of respondents (65 of 68) answered this question correctly.

- Cluster analysis 1
- Relevance analysis 1
- Anomaly analysis 65 ✓
- Regression analysis 1



9. **Data cleaning involves filling in missing values.** (1 point)

88% of respondents (60 of 68) answered this question correctly.

- True 60 ✓
- False 8



10. **Data integration is not a step in data preprocessing.** (1 point)

76% of respondents (52 of 68) answered this question correctly.

True	16	
False	52	✓



11. **Which of the following is not a form of data transformation?** (1 point)

76% of respondents (52 of 68) answered this question correctly.

Normalization	8	
Discretization	0	
Concept hierarchy	8	
Compression	52	✓



12. **In the diagram given below, depicting the KDD process, what is the correct order of tasks performed?** (1 point)

99% of respondents (67 of 68) answered this question correctly.

Selection, Preprocessing, Tran...	67	✓
Selection, Preprocessing, Tran...	0	
Interpretation, Preprocessing, ...	1	
Data mining, Preprocessing, Tr...	0	



13. **Sometimes the data under analysis has too many attributes, some of which may not be useful as per the task being performed. This problem is referred to as ____** (1 point)

93% of respondents (63 of 68) answered this question correctly.

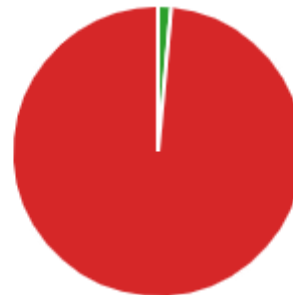
Overfitting	2	
Dimensionality curse	63	✓
Outliers in data	0	
Missing data	3	



14. **What is the full form of KDD in the data mining process?** (1 point)

99% of respondents (67 of 68) answered this question correctly.

Knowledge data house	0	
Knowledge data definition	0	
Knowledge discovery data	1	
Knowledge discovery database	67	✓



15. **Among which of the following can be used by the warehouse?** (1 point)

85% of respondents (58 of 68) answered this question correctly.

Database table	9	
Online database	0	
Flat files	1	
All of the above	58	✓



16. **Total fact table the star schema is composed of?** (1 point)

96% of respondents (65 of 68) answered this question correctly.

4	0
3	3
2	0
1	65 ✓



17. _____ **is the heart of the warehouse** (1 point)

100% of respondents (68 of 68) answered this question correctly.

data mining database servers.	0
data warehouse database serv...	68 ✓
data mart database servers	0
relational data base servers	0



18. **Which clustering algorithm does not require specifying the number of clusters in advance?** (1 point)

93% of respondents (63 of 68) answered this question correctly.

K-Means	3
DBSCAN	63 ✓
Agglomerative	2
Mean-Shift	0



19. _____ is a good alternative to the star schema.

(1 point)

65% of respondents (44 of 68) answered this question correctly.

● star schema.	0
● snowflake schema	23
● fact constellation	44 ✓
● star-snowflake schema	1



20. **Fact tables are** _____. (1 point)

32% of respondents (22 of 68) answered this question correctly.

● completely demoralized.	44
● partially demoralized.	1
● completely normalized.	22 ✓
● partially normalized.	1



21. **Which of the following is a goal of clustering algorithms?** (1 point)





91% of respondents (62 of 68) answered this question correctly.

● Classification	5
● Regression	0
● Dimensionality reduction	1
● Grouping similar data points t...	62 ✓



22. **Which clustering algorithm is based on the concept of centroids?** (1 point)

99% of respondents (67 of 68) answered this question correctly.

 K-Means	67	✓
 DBSCAN	0	
 Agglomerative	1	
 Mean-Shift	0	



DATA MINING AND DATA WAREHOUSING

9 Jul 2024

Totally 20 Marks

Total: 20 Points

🕒 20 minutes

* Required

1

USN *

2

NAME *

3

What is data mining * (1 Point)

- Deleting unnecessary data
- Sorting data alphabetically
- Storing data securely
- Extracting useful patterns or information from large datasets

4

Which of the following is not a basic data mining task? * (1 Point)

- Spooling
- Prediction
- Classification
- Clustering

5

Which of the following is not an issue in data mining? * (1 Point)

- High dimensionality
- Shortage of data
- Overfitting
- Outliers

6

Which among the following are not among Various Operations in OLAP Cube? * (1 Point)

- Sticking
- Dice
- Drill down
- Roll up

7

Which of the following is true about bottom up discretization? *

(1 Point)

- Some the values are treated as potential split points
- All the values are treated as potential split points
- Split points are not considered
- Only one value are treated as potential split points

8

The analysis of the data points that deviate from the general expected behavior of data in the data set is called _____ * (1 Point)

- Cluster analysis
- Relevance analysis
- Anomaly analysis
- Regression analysis

9

Data cleaning involves filling in missing values. * (1 Point)

- True
- False

10

Data integration is not a step in data preprocessing. * (1 Point)

- True
- False

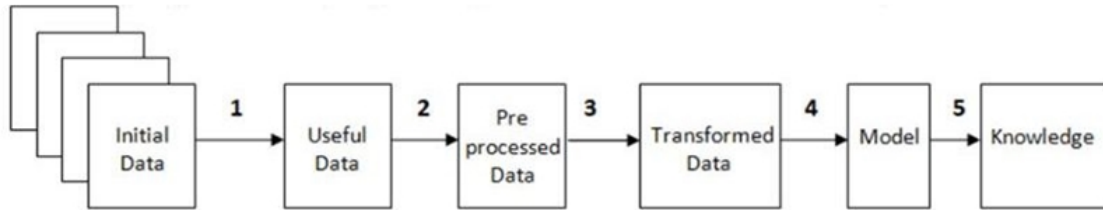
11

Which of the following is not a form of data transformation? *
(1 Point)

- Normalization
- Discretization
- Concept hierarchy
- Compression

12

In the diagram given below, depicting the KDD process, what is the correct order of tasks performed? * (1 Point)



- Selection, Preprocessing, Transformation, Data mining, Interpretation
- Selection, Preprocessing, Transformation, Interpretation, Data mining
- Interpretation, Preprocessing, Transformation, Data mining, Selection
- Data mining, Preprocessing, Transformation, Selection, Interpretation

13

Sometimes the data under analysis has too many attributes, some of which may not be useful as per the task being performed. This problem is referred to as _____ * (1 Point)

- Overfitting
- Dimensionality curse
- Outliers in data
- Missing data

14

What is the full form of KDD in the data mining process? * (1 Point)

- Knowledge data house
- Knowledge data definition
- Knowledge discovery data
- Knowledge discovery database

15

Among which of the following can be used by the warehouse? *
(1 Point)

- Database table
- Online database
- Flat files
- All of the above

16

Total fact table the star schema is composed of? * (1 Point)

- 4
- 3
- 2
- 1

17

 is the heart of the warehouse
* (1 Point)

- data mining database servers.
- data warehouse database servers
- data mart database servers
- relational data base servers

18

Which clustering algorithm does not require specifying the number of clusters in advance? * (1 Point)

- K-Means
- DBSCAN
- Agglomerative
- Mean-Shift

19

_____ is a good alternative to the star schema. * (1 Point)

- star schema.
- snowflake schema
- fact constellation
- star-snowflake schema

20

Fact tables are _____. * (1 Point)

- completely demoralized.
- partially demoralized.
- completely normalized.
- partially normalized.

21

Which of the following is a goal of clustering algorithms? * (1 Point)

- Classification
- Regression
- Dimensionality reduction
- Grouping similar data points together

Which clustering algorithm is based on the concept of centroids? *
(1 Point)

- K-Means
- DBSCAN
- Agglomerative
- Mean-Shift

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.