

{NO TITLE FOUND}

MULTIPLE CHOICE

1. What is KDD in data mining?

- | | |
|--------------------------------|----------------------------|
| a Knowledge Discovery Database | c Knowledge Discovery Data |
| b Knowledge Data definition | d Knowledge data house |

ANS: A PTS: 1

2. A _____ is a repository of information collected from multiple squares stored under a unified schema, and which usually resides at a single site.

- | | |
|---------------|--------------------|
| a Data mining | c Data warehouse |
| b Database | d legacy databases |

ANS: C PTS: 1

3. Which of the following databases is used to store time-related data?

- | | |
|---------------------|------------------------|
| a Spatial databases | c Multimedia databases |
| b Text databases | d Temporal databases |

ANS: D PTS: 1

4. Which of the following is the most popularly available and rich information repo

sitories?

- | | |
|------------------------|---------------------------|
| a Temporal databases | c Transactional databases |
| b Relational databases | d spatial databases |

ANS: B PTS: 1

5. A data normalization technique for real-valued attributes that divides each nume

rical

value by the same power of 10

- | | |
|-------------------------|-------------------|
| a Min-Max Normalization | c Decimal Scaling |
|-------------------------|-------------------|



Edit with WPS Office

- .
b Z-Score Normalization .
d Decimal Smoothing
.

ANS: C PTS: 1

6. A data warehouse is which of the following?

- a Can be updated by end users
. b Contains numerous naming conventions and formats
. c Organized around important subject areas
. d Contains only current data
.

ANS: C PTS: 1

7. The data Warehouse is_____

- a read only
. b write only
. c read and write only
. d none
.

ANS: A PTS: 1

8. Data about data is called

- a Table . c Metadata
. b Database . d Integration
.

ANS: C PTS: 1

9. _____ is the heart of the warehouse.

- a Data mining database servers. . c Data warehouse database servers.
. b Data mart database servers . d Relational data base servers.
.

ANS: C PTS: 1

10. Data Mining is also referred to as

- a Knowledge discovery in . c Data Extraction
. databases .



Edit with WPS Office

- ANS: A PTS: 1

0
decrease data storage requirements

- ANS: D PTS: 1

- a To explain some observed event or condition
- b To confirm that data exists
- c To analyze data for expected relationships
- d To create a new data warehouse

ANS: C PTS: 1

- a Historical data
- b Live data
- c Future Data
- d All of the above

ANS: A PTS: 1

- a The Waterfall Model
- b Object-Oriented Programming
- c The Scientific Method
- d Procedural Model

ANS: C PTS: 1

- A process to change data from a summary level to a detailed level
- A process to change data from a detailed level to a summary level



- c Joining data from one source into various sources of data
- d Separating data from one source into various sources of data

ANS: B PTS: 1

16. Data mining is best described as the process of _____

- a identifying patterns in data
- b deducing relationships in data
- c representing data
- d simulating trends in data

ANS: A PTS: 1

17. Reconciled data is which of the following?

- a Data stored in the various operational systems throughout the organization
- b Current data intended to be the single source for all decision support systems
- c Data stored in one operational system
- d Data that has been selected and formatted for end-user support applications

ANS: B PTS: 1

18. The _____ operation performs a selection on one dimension of the given cube, resulting in a sub cube

- a Pivot
- b Slice
- c Roll-up
- d Drilldown

ANS: B PTS: 1

19. OLTP systems are used for

- a Access to operational and historical data
- b Providing various add-ons for quick decision making
- c Supporting day-to-day business process
- d Providing complicated computations



Edit with WPS Office

ANS: C PTS: 1

20. OLAP stands for

- | | | | |
|---|-----------------------------------|---|--------------------------------|
| a | On-Line Analytical Processing | c | On-Line Access Processing |
| b | Operational Analytical Processing | d | On-Line Analytical Programming |

ANS: A PTS: 1

21. This data transformation technique works well when minimum and maximum values fo

r
a real-valued attribute are known

- | | | | |
|---|-----------------------|---|---------------------------|
| a | Min-Max Normalization | c | Z-Score Normalization |
| b | Decimal Scaling | d | Logarithmic Normalization |

ANS: A PTS: 1

22. This process removes redundancies that may be present in a data model

- | | | | |
|---|-----------------|---|-----------------|
| a | Abstraction | c | Standardization |
| b | Granularization | d | Normalization |

ANS: D PTS: 1

23. The set of values allowed to be entered into a column is known as its

- | | | | |
|---|-------------|---|-------------------|
| a | Data type | c | Domain |
| b | Constraints | d | None of the above |

ANS: C PTS: 1

24. The relational database model is designed to

- | | |
|---|---|
| a | Promote Data Redundancy |
| b | Minimize Data Redundancy |
| c | Eliminate the Need for Data Transformations |
| d | Eliminate the Need for Data Preprocessing |

ANS: B PTS: 1



Edit with WPS Office

25. Which of the following process includes data cleaning, data integration, data se

lection,

data transformation, data mining, and pattern evolution and knowledge presentati

on?

a KDD process

c KTL process

b ETL process

d None of the above

ANS: A PTS: 1

26. Which of the following Variable measured or their values defined, using one of t

wo

kinds of Nonmetric scales

a Numerical variable

c Categorical Variable

b Ordinal Variable

d Discrete Variable

ANS: D PTS: 1

27. Which of the following is not a characteristic of a data warehouse?

a Contains Nonvolatile Data

b Is Subject Oriented

c Supports Data Processing, Collection and Management

d Stores Data to be Reported on, Analyzed and Tested

ANS: C PTS: 1

28. This step of the KDD process model deals with noisy data

a Creating a Target Dataset

c Data Transformation

b Data Preprocessing

d Data Mining

ANS: B PTS: 1

29. The derived model may be represented in the form of



Edit with WPS Office

- a ER model
- b Flow chart
- c Decision trees
- d DFD

ANS: C PTS: 1

30. Which of the following is the classification of data mining systems?

- a Summarization
- b Visualization
- c Discrimination
- d Characterization

ANS: B PTS: 1

31. What are the functions of Data Mining??

- a Association and correctional analysis classification
- b Prediction and characterization
- c Cluster analysis and Evolution analysis
- d All of the above

ANS: D PTS: 1

32. Which of the following refers to the steps of the knowledge discovery process, i

n which the several data sources are combined??

- a Data selection
- b Data transformation
- c Data cleaning
- d Data integration

ANS: D PTS: 1

33. The data is stored, retrieved & updated in

- a OLAP
- b OLTP
- c SMTP
- d FTP

ANS: B PTS: 1

34. Expansion for DSS in Data Warahouse is?

- a Decision Support system.



Edit with WPS Office

- b Decision Single System
- c Data Storable System
- d Data Support System

ANS: A PTS: 1

35. _____ predicts future trends & behaviors, allowing business managers to make proactive, knowledge-driven decisions.

- a Data warehouse
- b Datamarts
- c Data mining
- d Metadata

ANS: C PTS: 1

36. Fact tables are?

- a completely demoralized
- b completely normalized.
- c partially demoralized
- d partially normalized

ANS: B PTS: 1

37. The star schema is composed of _____ fact table

- a one
- b two
- c three
- d four

ANS: A PTS: 1

38. An ___ system focuses mainly on the current data with in an enterprise or department, without referring to historical data or data in different organizations

- a On-Line Analytical Processing
- b On-Line Data Processing
- c On-Line Electronic Processing
- d On-Line Transaction Processing

ANS: D PTS: 1

39. A DWH is a subject oriented, integrated, time-variant, and _____ collection of data in support of management's decision-making process



Edit with WPS Office

- a Nonvolatile
- b Volatile
- c Disintegrated
- d Object- oriented

ANS: A PTS: 1

40. Data can be updated in ____environment.

- a data warehouse
- b data mining
- c operational
- d informational

ANS: C PTS: 1

41. Choose the _____ that will populate each fact table record

- a Measures
- b Dimensions
- c Grain
- d Business Process

ANS: A PTS: 1

42. Meta data repository contains

- a Operational meta data
- b Data irrelevant to system performance
- c The mapping from the DWH to the operational environment
- d Summarized data

ANS: A PTS: 1

43. A _____ contains a subset of corporate-wide data that is of value to a specifi

- c group of users.
- a Enterprise warehouse
- b Virtual warehouse
- c Data warehouse
- d Data mart

ANS: D PTS: 1

44. _____approach rules in rules are descending order of their priority

- a Ordered rules
- b Unordered rules
- c Rule set
- d None



Edit with WPS Office

ANS: A PTS: 1

45. _____are greedy algorithms for the induction of decision trees

- | | |
|--------|-------------------|
| a ID3 | c Both A and B |
| b C4.5 | d Backpropagation |

ANS: C PTS: 1

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

- | | |
|--------------------|-------------------------|
| a Star schema | c Star-snowflake schema |
| b Snowflake schema | d Fact constellation |

ANS: D PTS: 1

47. _____ is a supervised learning

- | | |
|---------------|------------------|
| a Cluster | c Classification |
| b Association | d prediction |

ANS: C PTS: 1

48. _____ the left-hand side of the rule is called

- | | |
|-------------------|------------|
| a Rule consequent | c Rule set |
| b Rule antecedent | d None |

ANS: B PTS: 1

49. _____ the right-hand side of the rule is called

- | | |
|-------------------|------------|
| a Rule consequent | c Rule set |
| b Rule antecedent | d None |

ANS: A PTS: 1

50. _____ is a unsupervised learning algorithm

- | | |
|---------------|------------------|
| a Cluster | c Classification |
| b Association | d prediction |

ANS: A PTS: 1



Edit with WPS Office

51. The type of relationship in star schema is _____

- | | | | |
|---|-------------|---|--------------|
| a | One-to-One | c | Many-to-One |
| b | One-to-Many | d | Many-to-Many |

ANS: A PTS: 1

52. Association rule Measured by

- a Support only
- b Confidence only
- c Support,Confidence only
- d Support,Confidence and lift only

ANS: D PTS: 1

53. In Hunt's algorithm , a decision tree is grown in _____ fashion

- | | | | |
|---|------------|---|-----------|
| a | Horizontal | c | Recursive |
| b | Vertical | d | None |

ANS: C PTS: 1

54. Nearest neighbor classifier makes their prediction based on _____ information

- | | | | |
|---|--------|---|-----------|
| a | Local | c | Temporary |
| b | Global | d | None |

ANS: A PTS: 1

55. _____ are responsible for running queries and reports against data warehouse tables.

- | | | | |
|---|-----------|---|-------------|
| a | Hardware. | c | End users |
| b | Software | d | Middle ware |

ANS: C PTS: 1

56. Prediction models



Edit with WPS Office

- | | | | |
|---|-----------------------------|---|------------|
| a | Classes | c | Attributes |
| . | | . | |
| b | Continuous-valued functions | d | None |
| . | | . | |

ANS: B PTS: 1

57. Classification is

- | | |
|---|---|
| a | A subdivision of a set of examples into a number of classes |
| . | |
| b | A measure of the accuracy, of the classification of a concept that is given |
| . | |
| c | by a certain theory |
| . | |
| d | The task of assigning a classification to a set of examples |
| . | |

ANS: A PTS: 1

58. _____ classification model can also be used to predict the class label of unk

nown records

- | | | | |
|---|----------------------|---|------------|
| a | Descriptive modeling | c | Regression |
| . | | . | |
| b | Predictive modelling | d | None |
| . | | . | |

ANS: B PTS: 1

59. _____classification model can serve as an explanatory toll to distinguish b

etween objects of different classes

- | | | | |
|---|----------------------|---|------------|
| a | Descriptive modeling | c | Regression |
| . | | . | |
| b | Predictive modelling | d | None |
| . | | . | |

ANS: A PTS: 1

60. In _____ approach the rules are in descending order of their priority

- | | | | |
|---|-----------------|---|----------|
| a | Ordered rules | c | Rule set |
| . | | . | |
| b | Unordered rules | d | None |
| . | | . | |

ANS: A PTS: 1

61. Query tool is meant for _____.



Edit with WPS Office

- a data acquisition
- b information delivery
- c information exchange.
- d communication

ANS: A PTS: 1

62. _____ Classification belongs to which type of learning?

- a supervised
- b unsupervised
- c rote
- d machine

ANS: A PTS: 1

63. _____ is a flow-chart like tree structure

- a Bayesion classifier
- b Decision tree
- c Backpropagaton
- d None

ANS: B PTS: 1

64. _____ is a eager learner classifier

- a Bayesion
- b Decision tree
- c Backpropagaton
- d Nearest neighbor

ANS: B PTS: 1

65. Accuracy can be defined as_____

- a Number of Correct predictions/Total number of predeictions
- b Number of wrong predictions/Total number of predeictions
- c Total number of predeictions/Total records
- d None

ANS: A PTS: 1

66. Classification predicts_____

- a Classes
- b Continuous-valued functions
- c Attributes
- d None

ANS: A PTS: 1



Edit with WPS Office

67. _____ not predicts the categorical labels

- | | | | |
|---|-------------|---|----------------|
| a | Cluster | c | Classification |
| . | | . | |
| b | Association | d | Prediction |
| . | | . | |

ANS: A PTS: 1

68. _____ node has no incomming edges

- | | | | |
|---|---------------|---|---------------|
| a | Internal node | c | root node |
| . | | . | |
| b | Leaf node | d | terminal node |
| . | | . | |

ANS: C PTS: 1

69. _____ predicts the categorical labels

- | | | | |
|---|-------------|---|----------------|
| a | Cluster | c | Classification |
| . | | . | |
| b | Association | d | Prediction |
| . | | . | |

ANS: C PTS: 1

70. _____ is a unsupervised learning algorithm

- | | | | |
|---|-------------|---|----------------|
| a | Cluster | c | Classification |
| . | | . | |
| b | Association | d | Prediction |
| . | | . | |

ANS: A PTS: 1

71. Data classification is _____ process

- | | | | |
|---|----------|---|------------|
| a | One-step | c | Three-step |
| . | | . | |
| b | Two-step | d | None |
| . | | . | |

ANS: B PTS: 1

72. Classification rules are extracted from _____

- | | |
|---|---------------|
| a | root node. |
| . | |
| b | siblings |
| . | |
| c | decision tree |
| . | |
| d | branches |
| . | |

ANS: C PTS: 1



73. Which of the following is constructed where the enterprise warehouse is the sole custodian of all warehouse data. Which is then distributed to the various depend

ent

data marts.

a Enterprise DWH

c Multi-tier DWH

b Two- tier DWH

d Virtual warehouse

ANS: C PTS: 1

74. Dimensionality reduction reduces the data set size by removing _____

a relevant attributes

c derived attributes.

b irrelevant attributes

d composite attributes.

ANS: B PTS: 1

75. A concept hierarchy that is a total or partial order among attributes in databases

is called a _____ hierarchy.

a Set-grouping

c Decision

b Grouping

d Schema

ANS: D PTS: 1

76. Classification accuracy is

a A subdivision of a set of examples into a number of classes

b No of examples correctly classified/total no of Examples

c The task of assigning a classification to a set of examples

d None of these

ANS: B PTS: 1

77. How many components are there in a data warehouse?

a two

c four



Edit with WPS Office

- b three
.
d five
.

ANS: D PTS: 1

78. The dimension tables describe the _____
a entities
.
b keys.
.
c facts.
.
d units of measures.
.

ANS: C PTS: 1

79. In Fact Table unique-id is referred as?
a primary key
.
b candidate key
.
c foreign key
.
d None
.

ANS: C PTS: 1

80. In Dimension table unique-id referred as?
a primary key
.
b candidate key
.
c foreign key
.
d none
.

ANS: A PTS: 1

81. Which of the following is the other name of Data mining?
a Exploratory data analysis.
.
b Data driven discovery
.
c Deductive learning.
.
d All of the above.
.

ANS: D PTS: 1

82. Converting data from different sources into a common format for processing is called as _____.
a selection.
.
b preprocessing.
.
c interpretation
.
d transformation.
.

ANS: D PTS: 1

83. Various visualization techniques are used in _____ step of KDD
a selection.
.
b transformation.
.
c data mining
.
d interpretation.
.

ANS: D PTS: 1

84. The proportion of transaction supporting X in T is called _____.
a confidence.
.
c support count.
.



- b support. d All of the above

ANS: B PTS: 1

85. The absolute number of transactions supporting X in T is called _____
a confidence c support count.
b support d None of the above

ANS: C PTS: 1

86. If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam, 10000 transaction contain both bread and jam. Then the support of bread and jam is _____
a 2% c 20%
b 3% d 30%

ANS: A PTS: 1

87. If T consist of 500000 transactions, 20000 transaction contain bread, 30000 transaction contain jam, 10000 transaction contain both bread and jam. Then the confidence of buying bread with jam is _____
a 33.33% c 45%
b 66.66% d 50%

ANS: D PTS: 1

88. Which of the following is not a frequent pattern mining algorithm?
a Apriori c Decission Tree
b Eclat d FP-Growth

ANS: C PTS: 1

89. Which algorithm requires fewer scans of data?
a FP-Growth c Apriori
b Eclat d Decission Tree

ANS: A PTS: 1

90. For the question given below consider the data Transactions :
1. I1, I2, I3, I4, I5, I6
2. I7, I2, I3, I4, I5, I6
3. I1, I8, I4, I5
4. I1, I9, I10, I4, I6
5. I10, I2, I4, I11, I5



With support as 0.6 find all frequent itemsets?

- | | | | |
|---|--|---|--|
| a | <I1>, <I2>, <I4>, <I5>, <I6>, <I1, I4>, <I2, I4>, <I2, I5>, <I4, I5>, <I4, I6>, <I2, I4, I5> | c | <I11>, <I4>, <I5>, <I6>, <I1, I4>, <I5, I4>, <I11, I5>, <I4, I6>, <I2, I4, I5> |
| b | <I2>, <I4>, <I5>, <I2, I4>, <I2, I5>, <I4, I5>, <I2, I4, I5> | d | <I1>, <I4>, <I5>, <I6> |

ANS: A PTS: 1

91. What techniques can be used to improve the efficiency of apriori algorithm?

- | | | | |
|---|-----------------------|---|------------|
| a | Hash-based techniques | c | Cleaning |
| b | Transaction Increases | d | prediction |

ANS: A PTS: 1

92. The first phase of Apriori algorithm is _____.

- | | | | |
|---|----------------------|---|--------------|
| a | Candidate generation | c | Pruning |
| b | Itemset generation | d | Partitioning |

ANS: A PTS: 1

93. The _____ step eliminates the extensions of (k-1)-itemsets which are not found to be frequent, from being considered for counting support

- | | | | |
|---|----------------------|---|------------|
| a | Candidate Generation | c | Pruning |
| b | Itemset Generation | d | Prediction |

ANS: C PTS: 1

94. The generalization of cross-tab which is represented visually is _____ which is also called as data cube.

- | | | | |
|---|-----------------------|---|--------------------|
| a | Two dimensional cube | c | N-dimensional cube |
| b | Multidimensional cube | d | Cuboid |

ANS: A PTS: 1

95. The operation of moving from finer-granularity data to a coarser granularity (by means of aggregation) is called a _____

- | | | | |
|---|------------|---|----------|
| a | Rollup | c | Dicing |
| b | Drill down | d | Pivoting |



Edit with WPS Office

ANS: A PTS: 1

96. _____ uses array-based multidimensional storage engines for multidimensional

views of data.

- | | |
|-----------|-----------|
| a KOLAP | c ZOLAP |
| · | · |
| b MOLAP | d ROLAP |
| · | · |

ANS: B PTS: 1

97. The pivot operation is also known as?

- | | |
|-----------------|----------|
| a Aggregation | c dice |
| · | · |
| b Rotation | d none |
| · | · |

ANS: B PTS: 1



Edit with WPS Office