

MARRUTHI SCHOOL OF BANKING

DATA MINING

1. is an essential process where intelligent methods are applied to extract data patterns.
A) Data warehousing B) Data mining C) Text mining D) Data selection
2. Data mining can also applied to other forms such as
i) Data streams ii) Sequence data iii) Networked data iv) Text data v) Spatial data
A) i, ii, iii and v only B) ii, iii, iv and v only C) i, iii, iv and v only D) All i, ii, iii, iv and v
3. Which of the following is not a data mining functionality?
A) Characterization and Discrimination B) Classification and regression
C) Selection and interpretation D) Clustering and Analysis
4. is a summarization of the general characteristics or features of a target class of data.
A) Data Characterization B) Data Classification C) Data discrimination D) Data selection
5. 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.
A) Data Characterization B) Data Classification C) Data discrimination D) Data selection
6. Strategic value of data mining is
A) cost-sensitive B) work-sensitive C) time-sensitive D) technical-sensitive
7. is the process of finding a model that describes and distinguishes data classes or concepts.
A) Data Characterization B) Data Classification C) Data discrimination D) Data selection
8. The various aspects of data mining methodologies is/are
i) Mining various and new kinds of knowledge ii) Mining knowledge in multidimensional space
iii) Pattern evaluation and pattern or constraint-guided mining. iv) Handling uncertainty, noise, or incompleteness of data
A) i, ii and iv only B) ii, iii and iv only C) i, ii and iii only D) All i, ii, iii and iv
9. The full form of KDD is
A) Knowledge Database B) Knowledge Discovery Database C) Knowledge Data House D) Knowledge Data Define
10. The out put of KDD is
A) Data B) Information C) Query D) Useful information
11. The full form of OLAP is
A) Online Analytical Processing B) Online Advanced Processing
C) Online Advanced Preparation D) Online Analytical Performance
12. is a subject-oriented, integrated, time-variant, nonvolatile collection or data in support of management decisions.
A) Data Mining B) Data Warehousing C) Document Mining D) Text Mining
13. The data is stored, retrieved and updated in
A) OLAP B) OLTP C) SMTP D) FTP
14. An system is market-oriented and is used for data analysis by knowledge workers, including managers, executives, and analysts.
A) OLAP B) OLTP C) Both of the above D) None of the above
15. is a good alternative to the star schema.
A) Star schema B) Snowflake schema C) Fact constellation D) Star-snowflake
16. The exposes the information being captured, stored, and managed by operational systems.
A) top-down view B) data warehouse view C) data source view D) business query view
17. The type of relationship in star schema is
A) many to many B) one to one C) one to many D) many to one

18. The allows the selection of the relevant information necessary for the data warehouse.
 A) top-down view B) data warehouse view C) data source view D) business query view
19. Which of the following is not a component of a data warehouse?
 A) Metadata B) Current detail data C) Lightly summarized data D) Component Key
20. Which of the following is not a kind of data warehouse application?
 A) Information processing B) Analytical processing C) Data mining D) Transaction processing
21. Data warehouse architecture is based on
 A) DBMS B) RDBMS C) Sybase D) SQL Server
22. supports basic OLAP operations, including slice and dice, drill-down, roll-up and pivoting.
 A) Information processing B) Analytical processing C) Data mining D) Transaction processing
23. The core of the multidimensional model is the, which consists of a large set of facts and a number of dimensions.
 A) Multidimensional cube B) Dimensions cube C) Data cube D) Data model
24. The data from the operational environment enter of data warehouse.
 A) Current detail data B) Older detail data C) Lightly Summarized data D) Highly summarized data
25. A data warehouse is
 A) updated by end users. B) contains numerous naming conventions and formats
 C) organized around important subject areas D) contain only current data
26. Business Intelligence and data warehousing is used for
 A) Forecasting B) Data Mining C) Analysis of large volumes of product sales data D) All of the above
27. Data warehouse contains data that is never found in the operational environment.
 A) normalized B) informational C) summary D) denormalized
28. are responsible for running queries and reports against data warehouse tables.
 A) Hardware B) Software C) End users D) Middle ware
29. The biggest drawback of the level indicator in the classic star schema is that it limits
 A) flexibility B) quantify C) qualify D) ability
30. are designed to overcome any limitations placed on the warehouse by the nature of the relational data model.
 A) Operational database B) Relational database C) Multidimensional database D) Data repository
31. 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 by a certain theory
 C. The task of assigning a classification to a set of examples D. None of these
32. Binary attribute are
 A. This takes only two values. In general, these values will be 0 and 1 and .they can be coded as one bit
 B. The natural environment of a certain species
 C. Systems that can be used without knowledge of internal operations D. None of these
33. Classification accuracy is
 A. A subdivision of a set of examples into a number of classes
 B. Measure of the accuracy, of the classification of a concept that is given by a certain theory
 C. The task of assigning a classification to a set of examples D. None of these

34. Classification task referred to
 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 by a certain theory
 C. The task of assigning a classification to a set of examples D. None of these
35. Cluster is
 A. Group of similar objects that differ significantly from other objects
 B. Operations on a database to transform or simplify data in order to prepare it for a machine-learning algorithm
 C. Symbolic representation of facts or ideas from which information can potentially be extracted D. None of these
36. Black boxes are
 A. This takes only two values. In general, these values will be 0 and 1 and they can be coded as one bit.
 B. The natural environment of a certain species
 C. Systems that can be used without knowledge of internal operations D. None of these
37. A definition of a concept is-----if it recognizes all the instances of that concept
 A. Complete B. Consistent C. Constant D. None of these
38. Data mining is
 A. The actual discovery phase of a knowledge discovery process
 B. The stage of selecting the right data for a KDD process
 C. A subject-oriented integrated time variant non-volatile collection of data in support of management D. Non
39. A definition or a concept is----- if it classifies any examples as coming within the concept
 A. Complete B. Consistent C. Constant D. None of these
40. Data selection is
 A. The actual discovery phase of a knowledge discovery process
 B. The stage of selecting the right data for a KDD process
 C. A subject-oriented integrated time variant non-volatile collection of data in support of management D. None
41. Classification task referred to
 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 by a certain theory
 C. The task of assigning a classification to a set of examples D. None of these
42. Discovery is
 A. It is hidden within a database and can only be recovered if one is given certain clues (an example IS encrypted information).
 B. The process of executing implicit previously unknown and potentially useful information from data
 C. An extremely complex molecule that occurs in human chromosomes and that carries genetic information in the form of genes. D. None of these
43. Data selection is
 A. The actual discovery phase of a knowledge discovery process
 B. The stage of selecting the right data for a KDD process
 C. A subject-oriented integrated time variant non-volatile collection of data in support of management D. None
44. Data scrubbing is which of the following?
 A. A process to reject data from the data warehouse and to create the necessary indexes
 B. A process to load the data in the data warehouse and to create the necessary indexes
 C. A process to upgrade the quality of data after it is moved into a data warehouse
 D. A process to upgrade the quality of data before it is moved into a data warehouse
45. The generic two-level data warehouse architecture includes which of the following?
 A. At least one data mart B. Data that can extracted from numerous internal and external sources
 C. Near real-time updates D. All of the above.

SOFTWARE ENGINEERING

- What is a DFD?
A. The modern version of flowchart
B. Mainly used at systems specification stage
C. The primary output of the system design phase
D. All of the above
- Checking quality of software in both simulated and live environments is known as
A. Checking
B. Usability
C. Validity
D. Validation
- The system should provide which thing, to avoid error in transcription and transposition, during data entry?
A. A check digit
B. A hand totals
C. Batch totals
D. All of these
- HIPO is
A. A forms driven technique in which standard forms are used to documents the information
B. Consists of a hierarchy chart and an associate set of input/process/output charts
C. Captures essence of top down decomposition
D. All of these
- DFD shows
A. The flow of data
B. The processes
C. The areas where they are stored
D. All of the above
- The time lag, between initiating a query and receiving a response is called
A. Response time
B. Waiting time
C. Processing time
D. Turn around time
- How are information system and capability related to each other ?
A. Improved processing speed
B. Increased volumed processing speed
C. Faster retrieval of information
D. All of the above
- The mistakes, committed by interchanging two digits in a numeic field, during data enty, is called
A. Transposition error
B. Transcription error
C. Beta testing error
D. Alpha testing error
- RAD Software process model stands for _____.
A) Rapid Application Development.
B) Relative Application Development.
C) Rapid Application Design.
D) Recent Application Development.
- Software project management comprises of a number of activities, which contains _____.
A. Project planning
B. Scope management
C. Project estimation
D. All mentioned above
- COCOMO stands for _____ Model.
A) Consumed COst Model
B) Constructive COst MOdel
C) Common COntrol Model
D) COmposition Cost
- Which of the following is not defined in a good Software Requirement Specification (SRS) document?
A) Functional Requirement.
B) Nonfunctional Requirement.
C) Goals of implementation.
D) Algorithm for software implementation.
- What is the simplest model of software development paradigm?
A) Spiral model
B) Big Bang model
C) V-model
D) Waterfall model
- Which of the following is not considered a tool at system design phase?
A. Data flow diagram
B. Decision table
C. Pie chart
D. System flow chart
- The best file organization in which one has to process all data records in a file called
A. Indexed
B. Sequential
C. Direct access
D. Random access
- Controls must be established in a system on order to
A. Prohibit tampering with information by unauthorized person
B. Verify that all data have been processed
C. The concurrent operation of an existing system and a new system
D. All of the above
- Coding and testing is done in following manner:
A. Adhoc
B. Cross sectional
C. Bottom-up
D. Top-clown
- The largest percentage of total life cycle cost of software is:
A. Design cost
B. Maintenance cost
C. Coding cost
D. Testing cost

20. In a system design project hardware design starts from.
A. Identification of size and capacity requirements of hardware B. Cataloging price from hardware vendors
C. Finding compatibility between hardware and software D. None of the above
21. Which of the following is not a major design consideration of the system.
A. Response time required B.Frequency of record updates C. Data integrity constant.
D.Availability of technically qualified personal to carry out design and development
22. Software maintenance includes
A. Bug fixing B.Setting preventing maintenance policy for servers
C.installation of software at site D. Designing of sotware for maintenance purposes.
23. Main diference between program testing and system testing is
A. Program testing is more comprehensive than system testing
B. System testing focuses on testing the interfaces between programs, program testing focuses on individual programs
C. System testing is tough and program testing is easy D. None of the above
24. The role of an integrated CASE tool in system analysis and design project is
A. Proving a CASE study for software development B. To provide tool for reacting a word file C.None of the above
D. To provide an environment that automate key tasks throughout sotware development process
25. The most creative and challenging phase of system life cycle is
A. Feasibility study B. Maintenance C. Design D. None of the above
26. The advantages of creating a prototype are.
A. It allows developers to experiment with number of diferent design options
B. It can serve as means of communication between developers and customers
C. Both (a) and (b) D. None of the above
- 27.The database design activity deals with the design of
A. Logical database B. Physical database C.Both (a) and (b) D.None of the above
28. From the following methods which size of the software product can be calculated?
A. Counting the lines of delivered code B. Counting delivered function points C. Both A and B D. None of the above
29. Which chart is a tool that depicts project as network diagram that is capable of graphically representing main events of project in both parallel and consecutive way?
A. PERT chart B. Gantt chart C. Both A & B D. None of the above
30. Coupling is a measure of
A. Relative functional strength B.Interdependence among module C.Both of the above D.None of the above