

STUDY MATERIALS

DSE-4, DATA MINING

(1)

VIth Semester

1) Data ware house is based on

- a) DOME
- b) ~~RDBMS~~
- c) sybase
- d) SQL server.

2) What is the process of data transformation?

- a) Comparison
- b) Projection
- c) Selection
- d) ~~Filtering~~

3. Which of the following is not involve in data mining?

- a) Knowledge extraction
- b) Data archaeology.
- c) Data exploration
- d) ~~Data transformation.~~

4) Relationships in data or patterns can be identified by.

- a) classification Model.
- b) Regression Model.
- c) descriptive Model.
- d) Predictive Model.

(2)

5) Which of the following maps data into predefined groups or classes.

- a) classification
- b) Regression
- c) clustering
- d) summarization

6) The problem of finding hidden structure in unlabeled data is called.

- a) supervised learning.
- b) unsupervised learning.
- c) Reinforcement learning
- d) serration.

7) Strategic value of data mining is .

- a) cost sensitive.
- b) work-sensitive.
- c) time-sensitive.
- d) technical sensitive.

8. Which of the following issue is considered before investing in data mining?

- a. Functionality
- b. Vendor consideration
- c. Compatibility
- d. all (a+b+c)

9. Querying of unstructured textual data is referred to as.

- a. Information access
- b. Information updation
- c. Information manipulation
- d. Information retrieval.

10. Correct process of Data Mining?

- a. Infrastructures, exploration, analysis, interpretation, exploitation.
- b. Exploitation, infrastructure, analysis, interpretation, Exploitation
- c. Exploitation, infrastructure, interpretation, analysis, exploitation.
- d. Exploration, infrastructures, analysis, exploitation, interpretation.

11. Data mining is

- a. Time variant non-volatile collection of data.
- b. Discovery phase of Knowledge.
- c. the stage of selecting the right data.
- d. none of the above.

12. The full form of KDD is. 4
- a. Knowledge Database
 - b. Knowledge data house.
 - c. Knowledge Discovery Database.
 - d. Knowledge Data Definition

13. Relationships ⁱⁿ data and patterns can be identified by.
- a. classification model.
 - b. Regression model.
 - c. clustering
 - d. Summarisation.

14.

13. which of the following divides data into groups that are meaningful, useful or both.
- a. cluster
 - b. Association
 - c. classification
 - d. Relation

14. Two fundamental goals of data mining are
- a. Analysis & description
 - b. Prediction & description
 - c. Data cleaning and organizing the data
 - d. Information & Knowledge

15. Key to represent relationship between tables is called.
- a. Primary Key
 - b. Secondary Key
 - c. Foreign Key
 - d. None of these.

16. Removing duplicate records in a data mining process called.

- a. data isolation b. Recovery c. data pruning
 d. data cleaning.

17. Analysis tools precompute summaries of very large amounts of data, in order to give

- a. queries response b. data access
c. Authorization d. consistency.

18. The output of KDD is

- a) data, b) information c) query
 d) useful information.

19. Which is a symbolic representation of facts from which information can potentially be extracted.?

- a. Knowledge b. data c. algorithm
d. program

20. Data visualization in mining cannot be done using

- a. Photos b. Graphs c. charts
d. Information Graphics.

21. Which of the following activities is a data mining task. 6.

- a. Monitoring the heart rate of a patient for abnormalities.
- b. Extracting the frequencies of a sound wave.
- c. Dividing the customers of a company according to their profitability.
- d. None of the above.

22. To detect fraudulent usage of credit cards, the following data mining task should be used.

- a. outlier analysis
- b. prediction
- c. association analysis
- d. feature selection.

23. Which data mining task can be used for predicting wind velocities as a function of temperature, humidity, air pressure, etc?

- a. cluster analysis
- b. Regression.
- c. classification
- d. sequential pattern discovery.

24. Synonym of data mining is

- a. Data warehouse
- b. Knowledge discovery in database.
- c. Business intelligence
- d. OLAP

25. Encountered or invalid data is known as.

- a. Missing data
- b. outlier
- c. Changing data
- d. noisy data.