

Name of subject: - Data Mining Techniques and Applications

Objective Type Question

1. What is true about data mining?

- A. Data Mining is defined as the procedure of extracting information from huge sets of data
- B. Data mining also involves other processes such as Data Cleaning, Data Integration, Data Transformation
- C. Data mining is the procedure of mining knowledge from data.
- D. All of the above

Answer - D

2. How many categories of functions involved in Data Mining?

- A. 2
- B. 3
- C. 4
- D. 5

Answer - D

3. The mapping or classification of a class with some predefined group or class is known as?

- A. Data Characterization
- B. Data Discrimination
- C. Data Set



D. Data Sub Structure

Ans : B

Explanation: Data Discrimination : It refers to the mapping or classification of a class with some predefined group or class

4. The analysis performed to uncover interesting statistical correlations between associated-attribute-value pairs is called?

A. Mining of Association

B. Mining of Clusters

C. Mining of Correlations

D. None of the above

Ans : C

Explanation: Mining of Correlations : It is a kind of additional analysis performed to uncover interesting statistical correlations between associated-attribute-value pairs or between two item sets to analyze that if they have positive, negative or no effect on each other.

5. _____ may be defined as the data objects that do not comply with the general behavior or model of the data available.

A. Outlier Analysis

B. Evolution Analysis

C. Prediction

D. Classification

Ans : A

Explanation: Outlier Analysis : Outliers may be defined as the data objects that do not comply with the general behavior or model of the data available.



6. "Efficiency and scalability of data mining algorithms" issues comes under?

- A. Mining Methodology and User Interaction Issues
- B. Performance Issues
- C. Diverse Data Types Issues
- D. None of the above

Ans : B

Explanation: In order to effectively extract the information from huge amount of data in databases, data mining algorithm must be efficient and scalable.

7. To integrate heterogeneous databases, how many approaches are there in Data Warehousing?

- A. 2
- B. 3
- C. 4
- D. 5

Ans : A

Explanation: Data warehousing involves data cleaning, data integration, and data consolidations. To integrate heterogeneous databases, we have the following two approaches : Query Driven Approach, Update Driven Approach.

8. Which of the following is correct advantage of Update-Driven Approach in Data Warehousing?

- A. This approach provides high performance.
- B. The data can be copied, processed, integrated, annotated, summarized and restructured in the semantic data store in advance.

C. Both A and B

D. None Of the above

Ans : C

Explanation: Both A and B are advantage of Update-Driven Approach in Data Warehousing.

9. What is the use of data cleaning?

A. to remove the noisy data

B. correct the inconsistencies in data

C. transformations to correct the wrong data.

D. All of the above

Ans : D

Explanation: Data cleaning is a technique that is applied to remove the noisy data and correct the inconsistencies in data. Data cleaning involves transformations to correct the wrong data. Data cleaning is performed as a data preprocessing step while preparing the data for a data warehouse.

10. Data Mining System Classification consists of?

A. Database Technology

B. Machine Learning

C. Information Science

D. All of the above

Ans : D

Explanation: A data mining system can be classified according to the following criteria : Database Technology, Statistics, Machine Learning, Information Science, Visualization, Other Disciplines

11. Which of the following is correct application of data mining?

A. Market Analysis and Management



B. Corporate Analysis & Risk Management

C. Fraud Detection

D. All of the above

Ans : D

Explanation: Data mining is highly useful in the following domains : Market Analysis and Management, Corporate Analysis & Risk Management, Fraud Detection

12. In Data Characterization, class under study is called as?

A. Study Class

B. Intial Class

C. Target Class

D. Final Class

Ans : C

Explanation: Data Characterization : This refers to summarizing data of class under study. This class under study is called as Target Class.

13. A sequence of patterns that occur frequently is known as?

A. Frequent Item Set

B. Frequent Subsequence

C. Frequent Sub Structure

D. All of the above

Ans : B

Explanation: Frequent Subsequence : A sequence of patterns that occur frequently such as purchasing a camera is followed by memory card.

14. _____ refers to the description and model regularities or trends for objects whose behavior changes over time.



- A. Outlier Analysis
- B. Evolution Analysis
- C. Prediction
- D. Classification

Ans : B

Explanation: Evolution Analysis : Evolution analysis refers to the description and model regularities or trends for objects whose behavior changes over time.

15. Pattern evaluation issue comes under?

- A. Mining Methodology and User Interaction Issues
- B. Performance Issues
- C. Diverse Data Types Issues
- D. None of the above

Ans : A

Explanation: Pattern evaluation : The patterns discovered should be interesting because either they represent common knowledge or lack novelty.

16. "Handling of relational and complex types of data" issue comes under?

- A. Mining Methodology and User Interaction Issues
- B. Performance Issues
- C. Diverse Data Types Issues
- D. None of the above

Ans : C

Explanation: The database may contain complex data objects, multimedia data objects, spatial data, temporal data etc. It is not possible for one



system to mine all these kind of data.

17. Which of the following is correct disadvantage of Query-Driven Approach in Data Warehousing?

- A. The Query Driven Approach needs complex integration and filtering processes.
- B. It is very inefficient and very expensive for frequent queries.
- C. This approach is expensive for queries that require aggregations.
- D. All of the above

Ans : D

Explanation: All statement are disadvantage of Query-Driven Approach in Data Warehousing.

18. The first steps involved in the knowledge discovery is?

- A. Data Integration
- B. Data Selection
- C. Data Transformation
- D. Data Cleaning

Ans : D

Explanation: The first steps involved in the knowledge discovery is Data Integration.

19. In which step of Knowledge Discovery, multiple data sources are combined?

- A. Data Cleaning
- B. Data Integration
- C. Data Selection
- D. Data Transformation

Ans : B



Explanation: Data Integration : multiple data sources are combined.

20. DMQL stands for?

- A. Data Mining Query Language
- B. Dataset Mining Query Language
- C. DBMiner Query Language
- D. Data Marts Query Language

Ans : A

Explanation: The Data Mining Query Language (DMQL) was proposed by Han, Fu, Wang, et al. for the DBMiner data mining system.