What does Apriori algorithm do?

a) **it mines all frequent patterns through pruning rules with lesser support**. c) Both a & b

b) it mines all frequent patterns through pruning rules with higher support.d) None of Above

Which of the following is an essential process in which the intelligent methods are applied to extract data patterns?

a) **Warehousing** b) Data Mining c)Text Mining d)Data Selection

What is KDD in data mining?

a) **Knowledge Discovery Database** b) Knowledge Discovery Data

c) Knowledge Data definition d) Knowledge data house

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

The data Warehouse is\_\_\_\_\_\_\_\_\_\_.

a) **Read only** b) write only c) read write only d) none

The data is stored, retrieved & updated in \_\_\_\_\_\_\_\_\_\_\_\_.

a) OLAP b) **OLTP** c) SMTP d) FTP

\_\_\_\_\_\_\_\_\_\_describes the data contained in the data warehouse.

a) Relational data b)operational data c) **Metadata** d)informational data.

The star schema is composed of \_\_\_\_\_\_\_\_\_\_ fact table.

a) **One** b) two c) three d) four.

Data can be updated in \_\_\_\_\_environment.

a) **Data warehouse** b) data mining c) operational d) informational

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is a data transformation process.

a) Comparison b) projection c) **selection** d) filtering.

\_\_\_\_\_\_\_\_\_\_\_\_ are some popular OLAP tools.

a) metacube, informix b) oracle express, essbase c) **holap** d) molap

Removing duplicate records is a process called \_\_\_\_\_\_\_\_\_\_\_\_\_.

a) Recovery b) **data cleaning** c) data cleansing d) data pruning

What is ETL?

a) Extract, Transform, Load b) Extract, Translate, Load

c) **Extract, Transfer, Load** d) Extract, Transcribe, Load

A \_\_\_\_\_\_\_\_\_\_\_ model identifies patterns or relationships.

a) Descriptive b) predictive c) **Regression** d) time series analysis

Which type of schema has dimensions with extended dimensions?

a) **Snow Flake** b) Star c) extended star d) Hybrid star

In \_\_\_\_\_\_\_\_ the groups are not predefined.

a) **Association rules** b) summarization c) clustering d) prediction

\_\_\_\_\_\_\_\_\_\_\_ is a good alternative to the star schema.

a) Star schema b) Snowflake schema c) **Fact constellation** d) Star-snowflake schema

Metadata contains at least \_\_\_\_\_\_\_\_\_.

a) **The structure of the data** b) the algorithms used for summarization

c) The mapping from the operational environment to the data warehouse

d) all of the above.

The dimension tables in data cubes describe the \_\_\_\_\_\_\_\_\_.

a) Entities b) **facts**  c) keys d) units of measures

The type of relationship in star schema is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

a) many-to-many b) one-to-one c) **one-to-many** d) many-to-one

Classification rules are extracted from \_\_\_\_\_\_\_\_\_\_\_\_\_.

a) Root node b) **Decision tree** c) siblings d) branches

Which of the following terms refers to a set of items that frequently occur together in a transactional database? A) **Frequent itemset** B) Classification C) Regression D) Clustering

Classification rules are extracted from \_\_\_\_\_\_\_\_\_\_\_\_\_.

a) Root node b) **Decision tree** c) siblings d) branches

|  |  |  |
| --- | --- | --- |
|  | | |
|  | | |
|  | Which of the following is an essential process in which the intelligent methods are applied to extract data patterns? a) **Warehousing** b) Data Mining c)Text Mining d)Data Selection |  |
|  | The data Warehouse is\_\_\_\_\_\_\_\_\_\_.  a) **Read only** b) write only c) read write only d) none |  |
|  | The data is stored, retrieved & updated in \_\_\_\_\_\_\_\_\_\_\_\_.  a) OLAP b) **OLTP** c) SMTP d) FTP |  |
|  | Data warehouse architecture is based on \_\_\_\_\_\_\_\_\_\_\_\_\_\_.  a) DBMS b) **RDBMS**  c) Sybase d) SQL Server. |  |
|  | Which of the following is not an issue in data mining?  a) High dimensionality b) **Shortage of data** c) Overfitting d) Outliers |  |
|  | Which of the following is a subset of data warehouse focused on a specific functional area?  a) **Data mart** b) Association rules c) Flat files d) Database |  |
|  | Which among the following are not among Various Operations in Data Warehousing?  a) **Sticking** b) Dice c) Drill down d) Roll up |  |
|  | \_\_\_\_\_\_\_\_\_\_\_ is a good alternative to the star schema.  a) Star schema b) Snowflake schema **c) Fact constellation** d)Star-snowflake schema |  |
|  | The type of relationship in star schema is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.  a) many-to-many b) one-to-one **c) one-to-many** d) many-to-one. |  |
|  | Classification rules are extracted from \_\_\_\_\_\_\_\_\_\_\_\_\_.  a) root node **b) decision tree** c) siblings d)branches |  |
|  | Which of the following is a drawback in manual filling of missing values? a) Incorrect data b) Redundancy in data c) Complex data d) **Too many missing values in the data** |  |
|  | Which of the following values are considered in binning for noise smoothing for a data value?  a) **Neighborhood**  b) Average c) Farthest d) Compressed |  |
|  | Which among the given is not among the basic statistical descriptions of data?  a) Mean b) Median c) **Average** d) Mode |  |
|  | A box plot displays the data distribution based on five-number summary.  a) **True** b) False |  |
|  | The type of relationship in star schema is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.  a) many-to-many b) one-to-one c) **one-to-many** d) many-to-one |  |
|  | Which type of schema has dimensions with extended dimensions?  a) **Snow Flake** b) Star c) extended star d) Hybrid star |  |
|  | What is the key concept behind the Apriori algorithm?  a) Using a divide-and-conquer approach to split the dataset  b) **Finding frequent itemsets by exploring subsets with a bottom-up approach**  c) Applying deep learning techniques to find patterns  d) Utilizing genetic algorithms for pattern mining |  |
|  | What is the primary goal of frequent pattern mining in data mining?  A) To compress data B) **To discover patterns and relationships in large datasets**  C) To sort and classify data D) To encrypt data |  |
|  | Classification rules are extracted from \_\_\_\_\_\_\_\_\_\_\_\_\_.  a) Root node b) **Decision tree** c) siblings d) branches |  |
|  | Which of the following terms refers to a set of items that frequently occur together in a transactional database? A) **Frequent itemset** B) Classification C) Regression D) Clustering |  |