

National Testing Agency

Question Paper Name: Data Mining 10th November 2019 Shift 1
Subject Name: Data Mining
Creation Date: 2019-11-10 13:22:32
Duration: 180
Total Marks: 100
Display Marks: Yes

Data Mining

Group Number : 1
Group Id : 709597115
Group Maximum Duration : 0
Group Minimum Duration : 120
Revisit allowed for view? : No
Revisit allowed for edit? : No
Break time: 0
Group Marks: 100

Data Mining

Section Id : 709597117
Section Number : 1
Section type : Online
Mandatory or Optional: Mandatory
Number of Questions: 100
Number of Questions to be attempted: 100
Section Marks: 100
Display Number Panel: Yes
Group All Questions: No

Sub-Section Number: 1
Sub-Section Id: 709597126
Question Shuffling Allowed : Yes

Question Number : 1 Question Id : 70959711372 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

A process known as _____ is used by large retailers to study trends.

- A. data mining
- B. data selection
- C. POS
- D. data conversion

Question Number : 2 Question Id : 70959711373 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

_____ is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management decisions.

- A. Data Mining.
- B. Data Warehousing.
- C. Web Mining.
- D. Text Mining.

Question Number : 3 Question Id : 70959711374 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The important aspect of the data warehouse environment is that data found within the data warehouse is _____.

- A. subject-oriented.
- B. time-variant.
- C. integrated.
- D. All of the above.

Question Number : 4 Question Id : 70959711375 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The time horizon in Data warehouse is usually _____.

- A. 1-2 years.
- B. 3-4years.
- C. 5-6 years.
- D. 5-10 years.

Question Number : 5 Question Id : 70959711376 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Data transformation includes which of the following?

- A. A process to change data from a detailed level to a summary level
- B. A process to change data from a summary level to a detailed level
- C. Joining data from one source into various sources of data
- D. Separating data from one source into various sources of data

Question Number : 6 Question Id : 70959711377 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following process includes data cleaning, data integration, data transformation, data selection, data mining, pattern evaluation and knowledge presentation?

- A. KDD process
- B. ETL process
- C. KTL process
- D. None of the above

Question Number : 7 Question Id : 70959711378 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Association rules are always defined on _____.

- A. Binary attribute.
- B. Single attribute.
- C. Relational database.
- D. Multidimensional attribute.

Question Number : 8 Question Id : 70959711379 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

_____ is data about data.

- A. Metadata.
- B. Microdata.
- C. Minidata
- D. Multidata.

Question Number : 9 Question Id : 70959711380 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Capability of data mining is to build _____ models.

- A. Retrospective.
- B. Interrogative.
- C. Predictive.
- D. Imperative.

Question Number : 10 Question Id : 70959711381 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.K-nearest neighbor is one of the _____.

- A. learning technique.
- B. OLAP tool.
- C. purest search technique.
- D. data warehousing tool

Question Number : 11 Question Id : 70959711382 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.In data mining Enrichment means _____.

- A. adding external data.
- B. deleting data.
- C. cleaning data.
- D. selecting the data.

Question Number : 12 Question Id : 70959711383 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following features usually applies to data in a data warehouse?

- A.Data are often deleted
- B. Most applications consist of transactions
- C.Data are rarely deleted
- D.Relatively few records are processed by applications

Question Number : 13 Question Id : 70959711384 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Which of the following statement is true?

- A. The data warehouse consists of data marts and operational data
- B. The data warehouse is used as a source for the operational data
- C. The operational data are used as a source for the data warehouse
- D. All of the above

Question Number : 14 Question Id : 70959711385 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.The following is true of three-tier data warehouses:

- A. Once created, the data marts will keep on being updated from the data warehouse at periodic times
- B. Once created, the data marts will directly receive their new data from the operational databases
- C. The data marts are different groups of tables in the data warehouse
- D. A data mart becomes a data warehouse when it reaches a critical size

Question Number : 15 Question Id : 70959711386 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The following technology is not well-suited for data mining:

- A. Expert system technology
- B. Data visualization
- C. Technology limited to specific data types such as numeric data types
- D. Parallel architecture

Question Number : 16 Question Id : 70959711387 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

What is true of the multidimensional model?

- A. It typically requires less disk storage
- B. It typically requires more disk storage
- C. Typical business queries requiring aggregate functions take more time
- D. Increasing the size of a dimension is difficult

Question Number : 17 Question Id : 70959711388 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Data transformation includes which of the following?

- A. A process to change data from a detailed level to a summary level
- B. A process to change data from a summary level to a detailed level
- C. Joining data from one source into various sources of data
- D. Separating data from one source into various sources of data

Question Number : 18 Question Id : 70959711389 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

A snowflake schema is which of the following types of tables?

- A. Fact
- B. Dimension
- C. Helper
- D. All of these

Question Number : 19 Question Id : 70959711390 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Fact tables are which of the following?

- A. Completely denormalized
- B. Partially denormalized
- C. Completely normalized
- D. Partially normalized

Question Number : 20 Question Id : 70959711391 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

A star schema has what type of relationship between a dimension and fact table?

- A. many to many
- B. One to one
- C. One to many
- D. All of the above

Question Number : 21 Question Id : 70959711392 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

What is ETL Stand for?

- A. Execute tramit and load
- B. Extract transform and load
- C. Excute Transform and load
- D. All the above

Question Number : 22 Question Id : 70959711393 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

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.

Question Number : 23 Question Id : 70959711394 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

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

Question Number : 24 Question Id : 70959711395 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The process of removing the deficiencies and loopholes in the data is called as

- A. Aggregation of data
- B. Extracting of data
- C. Cleaning up of data.
- D. Loading of data

Question Number : 25 Question Id : 70959711396 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is not related to dimension table attributes?

- A. Descriptive
- B. Equally unavailable
- C. Complete
- D. Indexed.

Question Number : 26 Question Id : 70959711397 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is the collection of data objects that are similar to one another within the same group?

- A. Partitioning
- B. Table
- C. Cluster
- D. Data source

Question Number : 27 Question Id : 70959711398 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

At which level we can create dimensional models?

- A. Business requirements level
- B. Architecture models level
- C. Implementation level
- D. Testing level.

Question Number : 28 Question Id : 70959711399 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one manages both current and historic transactions?

- A. OLTP
- B. OLAP
- C. Spread sheet
- D. XML

Question Number : 29 Question Id : 70959711400 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is the most important when deciding on the data structure of a data mart?

- A. Metadata naming conventions
- B. Data access tools to be used
- C. Extract, Transform, and Load (ETL) tool to be used
- D. All of these

Question Number : 30 Question Id : 70959711401 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

..... 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

Question Number : 31 Question Id : 70959711402 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

..... 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

Question Number : 32 Question Id : 70959711403 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Strategic value of data mining is

- A) cost-sensitive
- B) work-sensitive
- C) time-sensitive
- D) technical-sensitive

Question Number : 33 Question Id : 70959711404 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

..... 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

Question Number : 34 Question Id : 70959711405 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

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

Question Number : 35 Question Id : 70959711406 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

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

- a) Two dimensional cube
- b) Multi-dimensional cube
- c) N-dimensional cube
- d) Cuboid

Question Number : 36 Question Id : 70959711407 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The process of viewing the cross-tab (Single dimensional) with a fixed value of one attribute is

- a) Slicing
- b) Dicing
- c) Pivoting
- d) Both Slicing and Dicing

Question Number : 37 Question Id : 70959711408 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In Data warehouse Record cannot be updated in _____.

- A. OLTP
- B. files
- C. RDBMS
- D. data warehouse

Question Number : 38 Question Id : 70959711409 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is a descriptive model?

- A. Classification.
- B. Regression.
- C. Sequence discovery.
- D. Association rules.

Question Number : 39 Question Id : 70959711410 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Knowledge is referred to

- A) Non-trivial extraction of implicit previously unknown and potentially useful information from data
- B) Set of columns in a database table that can be used to identify each record within this table uniquely
- C) Collection of interesting and useful patterns in a database
- D) none of these

Question Number : 40 Question Id : 70959711411 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

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

Question Number : 41 Question Id : 70959711412 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Enrichment means

- A. Adding external data
- B. Deleting data
- C. Cleaning Data
- D. Selecting the data

Question Number : 42 Question Id : 70959711413 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The decision support system is used only for

- A. Cleaning
- B. Coding
- C. Selecting
- D. queries

Question Number : 43 Question Id : 70959711414 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Deep knowledge can be found only by using

- A. Clues
- B. OLAP
- C. SQL
- D. Algorithm

Question Number : 44 Question Id : 70959711415 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Metadata describes

- A. Content of database
- B. Structure of content of database
- C. structure of database
- D. Database itself

Question Number : 45 Question Id : 70959711416 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The partition of overall data warehouse is

- A. Database
- B. Data cube
- C. Data mart
- D. Operational data

Question Number : 46 Question Id : 70959711417 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one is an example for case based learning

- A. Decision tree
- B. Neural network
- C. Genetic algorithm
- D. K-Nearest neighbor

Question Number : 47 Question Id : 70959711418 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

k-nearest neighbour is one of the

- A. Learning technique
- B. OLAP tool
- C. Purest search technique
- D. Data warehousing tool

Question Number : 48 Question Id : 70959711419 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The complexity of data mining algorithm is represented by

- A. $\log n$
- B. $2n \log n$
- C. $n \log n$
- D. $2 \log n$

Question Number : 49 Question Id : 70959711420 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The intermediate unit in perception is

- A. Photoreceptors
- B. Associators
- C. Responders
- D. Receptors

Question Number : 50 Question Id : 70959711421 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In KDD and data mining noise is referred to as

- A. Repeated data
- B. Complex data
- C. Meta data
- D. Random errors in database

Question Number : 51 Question Id : 70959711422 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

With data mining, the best way to accomplish this is by setting aside some of your data in a vault to isolate it from the mining process; once the mining is complete, the results can be tested against the isolated data to confirm the model's _____.

- A. Validity
- B. Security
- C. Integrity
- D. None of above

Question Number : 52 Question Id : 70959711423 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Classification consists of examining the properties of a newly presented observation and assigning it to a predefined _____.

- A.Object
- B.Container
- C.Subject
- D.Class

Question Number : 53 Question Id : 70959711424 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

During business hours, most _____ systems should probably not use parallel execution.

- A.OLAP
- B.DSS
- C.Data Mining
- D.OLTP

Question Number : 54 Question Id : 70959711425 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The users of data warehouse are knowledge workers in other words they are _____ in the organization.

- A.Decision maker
- B.Manager
- C.Database Administrator
- D.DWH Analyst

Question Number : 55 Question Id : 70959711426 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Multi-dimensional databases (MDDs) typically use _____ formats to store pre-summarized cube structures.

- A.SQL
- B.proprietary file
- C.Object oriented
- D.Non- proprietary file

Question Number : 56 Question Id : 70959711427 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Data warehousing and on-line analytical processing (OLAP) are _____ elements of decision support system.

- A.Unusual
- B.Essential
- C.Optional
- D.None of the given

Question Number : 57 Question Id : 70959711428 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Analytical processing uses _____, instead of record level access.

- A.multi-level aggregates
- B.Single-level aggregates
- C.Single-level hierarchy
- D.None of the Given

Question Number : 58 Question Id : 70959711429 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The divide & conquer cube partitioning approach helps alleviate the _____ limitations of MOLAP implementation.

- A.Flexibility
- B.Maintainability
- C.Security
- D.Scalability

Question Number : 59 Question Id : 70959711430 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Data Warehouse provides the best support for analysis while OLAP carries out the _____ task.

- A.Mandatory
- B.Whole
- C.Analysis
- D.Prediction

Question Number : 60 Question Id : 70959711431 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Data Warehouse provides the best support for analysis while OLAP carries out the _____ task.

- A.Mandatory
- B.Whole
- C.Analysis
- D.Prediction

Question Number : 61 Question Id : 70959711432 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In contrast to statistics, data mining is _____ driven.

- A.Assumption
- B.Knowledge
- C.Human
- D.Database

Question Number : 62 Question Id : 70959711433 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

As opposed to the outcome of classification, estimation deal with _____ valued outcome.

- A.Discrete
- B.Isolated
- C.Continuous
- D.Distinct

Question Number : 63 Question Id : 70959711434 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The goal of ideal parallel execution is to completely parallelize those parts of a computation that are not constrained by data dependencies. The smaller the portion of the program that must be executed _____, the greater the scalability of the computation.

- A.In Parallel
- B.Distributed
- C.Sequentially
- D.None of above

Question Number : 64 Question Id : 70959711435 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The most important part of _____ is selecting the variables on which clustering is based. A.interpreting and profiling clusters
B.selecting a clustering procedure
C.assessing the validity of clustering
D.formulating the clustering problem

Question Number : 65 Question Id : 70959711436 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

_____ is a clustering procedure characterized by the development of a tree-like structure.

- A.Non-hierarchical clustering
- B.Hierarchical clustering
- C.Divisive clustering
- D.Agglomerative clustering

Question Number : 66 Question Id : 70959711437 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Challenges of clustering includes?

- A. High dimensionality of data
- B. Scalability
- C. Noisy data
- D. All of the above

Question Number : 67 Question Id : 70959711438 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which is not a type of clustering?

- A. Density based
- B. Similarity based
- C. Decision driven
- D. Partition Based

Question Number : 68 Question Id : 70959711439 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is not clustering technique?

- A. Agglomerative
- B. Derivative
- C. Partitioning
- D. Density Based

Question Number : 69 Question Id : 70959711440 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In the K-means clustering algorithm the distance between cluster centroid to each object is calculated using _____method.

- A. Cluster distance
- B. Euclidean distance
- C. Cluster width
- D. None of above

Question Number : 70 Question Id : 70959711441 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

k-means clustering is also referred to as?

- A. Non-hierarchical clustering
- B. Optimizing partitioning
- C. Divisive clustering
- D. Agglomerative clustering

Question Number : 71 Question Id : 70959711442 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.The process of forming general concept definitions from examples of concept to be learned

- A.Deduction
- B.Abduction
- C.Induction
- D.Conjunction

Question Number : 72 Question Id : 70959711443 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Data mining is best described as the process of

- A. Identifying patterns data
- B. Deducting relationships in data
- C. Representing data
- D. simulating trends in data

Question Number : 73 Question Id : 70959711444 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Supervised learning and unsupervised clustering both require at least one

- A. Hidden attribute
- B. Input/output attribute
- C. Categorical attribute
- D. All of these

Question Number : 74 Question Id : 70959711445 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.At which level we can create dimensional models?

- A. Detailed models level
- B. Architecture models level
- C. Implementation level
- D. Testing level.

Question Number : 75 Question Id : 70959711446 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Which of the following is the collection of data objects that are similar to one another within the same group?

- A. Partitioning
- B. Cluster
- C. Table
- D. Data source.

Question Number : 76 Question Id : 70959711447 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Which of the following is not an ETL tool?

- A. DT/studio.
- B. Oracle warehouse builder
- C. Datastage
- D. Visual studio

Question Number : 77 Question Id : 70959711448 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.What is/are the different types of Meta data?

- I.Administrative
 - II.Business
 - III.Operational.
- A.Only I
B.Both (I) and (II) above
C.Both (II) and (III) above
D.All (I), (II) and (III) above

Question Number : 78 Question Id : 70959711449 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The apriori property means

- A.To improve the efficiency the level-wise generation of frequent item sets
- B.If a set can pass a test, all of its supersets will fail the same test as well
- C.To decrease the efficiency the level-wise generation of frequent item sets
- D.All of these

Question Number : 79 Question Id : 70959711450 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

An OLAP tool provides for

- A.Multidimensional analysis
- B.Roll-up and drill-down
- C.Slicing and dicing
- D.Rotation

Question Number : 80 Question Id : 70959711451 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Data modeling technique used for data marts is

- A.Dimensiona modeling
- B.ER – model
- C.Extended ER – model
- D.Logical model.

Question Number : 81 Question Id : 70959711452 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Business Intelligence (BI) reporting analyses can be performed using _____ .

- A.standard SQL only
- B.extensions to SQL only
- C.OLAP only
- D.both standard SQL and extensions to SQL

Question Number : 82 Question Id : 70959711453 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

.Data mining applications are used accomplish which of the following tasks?

- A.Process transactions only
- B.Do RFM analysis only
- C.Do what-if analysis only
- D.Do both RFM and what-if analysis

Question Number : 83 Question Id : 70959711454 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which of the following is an unsupervised data mining technique?

- A.Cluster analysis only
- B.Regression Analysis only
- C.RFM Analysis only
- D.Both Regression Analysis and RFM Analysis

Question Number : 84 Question Id : 70959711455 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In OLAP, we are interested in _____

- A.levels only
- B.dimensions only
- C.measures only
- D.measures and dimensions

Question Number : 85 Question Id : 70959711456 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

We have Market Basket data for 1,000 rental transactions at a Video Store. There are four videos for rent -- Video A, Video B, Video C and Video D. The probability that Video D will be rented given that Video C has been rented is known as _____

- A.the basic probability
- B.support
- C.lift
- D.confidence

Question Number : 86 Question Id : 70959711457 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

One popular classification technique in Business Intelligence (BI) reporting is

- A.cluster analysis only
- B.regression analysis only
- C.RFM analysis only
- D.both cluster analysis and regression analysis

Question Number : 87 Question Id : 70959711458 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which operation mainly used to access the live data online and analyse it

- A.ROLAP
- B.OLAP
- C.MOLAP
- D.HOLAP

Question Number : 88 Question Id : 70959711459 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In data mining, the slice operation performs a selection on

- A.One dimension
- B.Two dimension
- C.Multi dimension
- D.None of the above

Question Number : 89 Question Id : 70959711460 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Preprocessing step is used in Data is

- A. cleaning
- B. integration
- C. transformation
- D. reduction

Question Number : 90 Question Id : 70959711461 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Total number of stages in KDD is

- A. 3
- B. 4
- C. 5
- D. 6

Question Number : 91 Question Id : 70959711462 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

The process of upgrading the existing quality of data before it is moved into a data warehouse is known as

- A. Data Scrubbing
- B. Data Polishing
- C. Data Auditing
- D. Data Aggregation

Question Number : 92 Question Id : 70959711463 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

In OLAP cubes, data (measures) are categorized by

- A. Length
- B. Angles
- C. Dimensions
- D. Breadth

Question Number : 93 Question Id : 70959711464 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

An OLAP system is market-oriented and is used for data analysis by knowledge workers, including managers, executives, and analysts.

- A. True
- B. False
- C. May be
- D. Cant say

Question Number : 94 Question Id : 70959711465 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one is not a Data Transformation strategy

- A. Aggregation
- B. Normalization
- C. Generalization
- D. Compression

Question Number : 95 Question Id : 70959711466 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

A subset of the data warehouse that contains all the data about a particular topic is known as

- A.Data Pocket
- B.Metadata
- C.Data Mart
- D.Data Mining

Question Number : 96 Question Id : 70959711467 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one is the last step of Data Mart Implementation process

- A.Designing
- B.Accessing
- C.Managing
- D.Populating

Question Number : 97 Question Id : 70959711468 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one is not a data reduction strategy

- A.Data Generalization
- B.Dimension reduction
- C.Data compression
- D.Data cube aggregation

Question Number : 98 Question Id : 70959711469 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which one is a data mining function that assigns items in a collection to target categories or classes

- A.Selection
- B.Classification
- C.Integration
- D.Reduction

Question Number : 99 Question Id : 70959711470 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Usually _____ years is the time horizon in data warehouse

- A.1-3
- B.3-5
- C.5-10
- D.10-15

Question Number : 100 Question Id : 70959711471 Question Type : MCQ Option Shuffling : No Display Question Number : Yes
Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Operational metadata defines the structure of the data held in operational databases and used by operational applications

- A.True
- B.Fasle
- C.May be
- D.Can't say