Data Mining

Multiple Choice Question & Answers:-

1) Adaptive system management is

A) It uses machine-learning techniques. Here program can learn from past experience and adapt themselves to new situations.

B) Computational procedure that takes some value as input and produces some value as output.

C) Science of making machines performs tasks that would require intelligence when performed by humans.

D) None of these

Answer: A

2) Bayesian classifiers is

A) A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.

B) Any mechanism employed by a learning system to constrain the search space of a hypothesis.

C) An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation.

D) None of these

Answer: A

3) Algorithm is

A) It uses machine-learning techniques. Here program can learn from past experience and adapt themselves to new situations.

B) Computational procedure that takes some value as input and produces some value as output.

C) Science of making machines performs tasks that would require intelligence when performed by humans.

D) None of these

Answer: B

4) Bias is

A) A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.

B) Any mechanism employed by a learning system to constrain the search space of a hypothesis.

C) An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation.

D) None of these

Answer: B

5) Background knowledge referred to

A) Additional acquaintance used by a learning algorithm to facilitate the learning process.

B) A neural network that makes use of a hidden layer.

C) It is a form of automatic learning.

D) None of these

Answer: A

6) Case-based learning is

A) A class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory.

B) Any mechanism employed by a learning system to constrain the search space of a hypothesis.

C) An approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation.

D) None of these

Answer: C

7) 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

Answer: A

8) 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

Answer: A

9) 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

Answer: B

10) Biotope 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

Answer: B

11) 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

Answer: A

12) 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

Answer: C

13) 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

Answer: A

14) 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) None of these

Answer: A

15) 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

Answer: B

16) 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 of these

Answer: B

17) 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

Answer: C

18) DNA (Deoxyribonucleic acid)

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

Answer: C

19) Hybrid is

A) Combining different types of method or information

B) Approach to the design of learning algorithms that is structured along the lines of the theory of evolution.

C) Decision support systems that contain an information base filled with the knowledge of an expert formulated in terms of if-then rules.

D) None of these

Answer: A

20) 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

Answer: B

21) Euclidean distance measure is

A) A stage of the KDD process in which new data is added to the existing selection.

B) The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them

C) The distance between two points as calculated using the Pythagoras theorem.

D) None of these

Answer: C

22) Hidden knowledge referred to

A) A set of databases from different vendors, possibly using different database paradigms

B) An approach to a problem that is not guaranteed to work but performs well in most cases

C) Information that is hidden in a database and that cannot be recovered by a simple SQL query.

D) None of these

Answer: C

23) Enrichment is

A) A stage of the KDD process in which new data is added to the existing selection

B) The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them

C) The distance between two points as calculated using the Pythagoras theorem.

D) None of these

Answer: A

24) Heterogeneous databases referred to

A) A set of databases from different b vendors, possibly using different database paradigms

B) An approach to a problem that is not guaranteed to work but performs well in most cases.

C) Information that is hidden in a database and that cannot be recovered by a simple SQL query.

D) None of these

Answer: A

25) Enumeration is referred to

A) A stage of the KDD process in which new data is added to the existing selection.

B) The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them

C) The distance between two points as calculated using the Pythagoras theorem.

D) None of these

Answer: B

26) Heuristic is

A) A set of databases from different vendors, possibly using different database paradigms

B) An approach to a problem that is not guaranteed to work but performs well in most cases

C) Information that is hidden in a database and that cannot be recovered by a simple SQL query.

D) None of these

Answer: B

27) Hybrid learning is

A) Machine-learning involving different techniques

B) The learning algorithmic analyzes the examples on a systematic basis 2nd makes incremental adjustments to the theory that is learned

C) Learning by generalizing from examples

D) None of these

Answer: A

28) Kohonen self-organizing map referred to

A) The process of finding the right formal representation of a certain body of knowledge in order to represent it in a knowledge-based system

B) It automatically maps an external signal space into a system's internal representational space. They are useful in the performance of classification tasks

C) A process where an individual learns how to carry out a certain task when making a transition from a situation in which the task cannot be carried out to a situation in which the same' task under the same circumstances can be carried out.

D) None of these

Answer: B

29) Incremental learning referred to

A) Machine-learning involving different techniques

B) The learning algorithmic analyzes the examples on a systematic basis and makes incremental adjustments to the theory that is learned

C) Learning by generalizing from examples

D) None of these

Answer: B

30) Knowledge engineering is

A) The process of finding the right formal representation of a certain body of knowledge in order to represent it in a knowledge-based system

B) It automatically maps an external signal space into a system's internal representational space. They are useful in the performance of classification tasks.

C) A process where an individual learns how to carry out a certain task when making a transition from a situation in which the task cannot be carried out to a situation in which the same task under the same circumstances can be carried out.

D) None of these

Answer: A

31) Information content is

A) The amount of information with in data as opposed to the amount of redundancy or noise.

B) One of the defining aspects of a data warehouse

C) Restriction that requires data in one column of a database table to the a subset of another-column.

D) None of these

Answer: A

32) Inductive learning is

A) Machine-learning involving different techniques

B) The learning algorithmic analyzes the examples on a systematic basis and makes incremental adjustments to the theory that is learned

C) Learning by generalizing from examples

D) None of these

Answer: C

33) Inclusion dependencies

A) The amount of information with in data as opposed to the amount of redundancy or noise

B) One of the defining aspects of a data warehouse

C) Restriction that requires data in one column of a database table to the a subset of another-column

D) None of these

Answer: C

34) KDD (Knowledge Discovery in Databases) 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

Answer: A

35) Learning is

A) The process of finding the right formal representation of a certain body of knowledge in order to represent it in a knowledge-based system

B) It automatically maps an external signal space into a system's internal representational space. They are useful in the performance of classification tasks.

C) A process where an individual learns how to carry out a certain task when making a transition from a situation in which the task cannot be carried out to a situation in which the same task under the same circumstances can be carried out.

D) None of these

Answer: C

36) Naive prediction is

A) A class of learning algorithms that try to derive a Prolog program from examples.

B) A table with n independent attributes can be seen as an n- dimensional space.

C) A prediction made using an extremely simple method, such as always predicting the same output.

D) None of these

Answer: C

37) Learning algorithm referrers to

A) An algorithm that can learn

B) A sub-discipline of computer science that deals with the design and implementation of learning algorithms.

C) A machine-learning approach that abstracts from the actual strategy of an individual algorithm and can therefore be applied to any other form of machine learning.

D) None of these

Answer: A

38) 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

Answer: C

39) Node is

 A) A component of a network

 B) In the context of KDD and data mining, this refers to random errors in a database table.

 C) One of the defining aspects of a data warehouse

 D) None of these

Answer: A

40) Machine learning is

 A) An algorithm that can learn

 B) A sub-discipline of computer science that deals with the design and implementation of learning algorithms

 C) An approach that abstracts from the actual strategy of an individual algorithm and can therefore be applied to any other form of machine learning.

 D) None of these

Answer: B

41) Projection pursuit is

 A) The result of the application of a theory or a rule in a specific case

 B) One of several possible enters within a database table that is chosen by the designer as the primary means of accessing the data in the table.

 C) Discipline in statistics that studies ways to find the most interesting projections of multi-dimensional spaces

 D) None of these

Answer: C

42) Inductive logic programming is

A) A class of learning algorithms that try to derive a Prolog program from examples

B) A table with n independent attributes can be seen as an n-dimensional space

C) A prediction made using an extremely simple method, such as always predicting the same output

D) None of these

Answer: A

43) Statistical significance is

A) The science of collecting, organizing, and applying numerical facts

B) Measure of the probability that a certain hypothesis is incorrect given certain observations.

C) One of the defining aspects of a data warehouse, which is specially built around all the existing applications of the operational data

D) None of these

Answer: B

44) Multi-dimensional knowledge is

A) A class of learning algorithms that try to derive a Prolog program from examples

B) A table with n independent attributes can be seen as an n-dimensional space

C) A prediction made using an extremely simple method, such as always predicting the same output.

D) None of these

Answer: B

45) Prediction is

A) The result of the application of a theory or a rule in a specific case

B) One of several possible enters within a database table that is chosen by the designer as the primary means of accessing the data in the table.

C) Discipline in statistics that studies ways to find the most interesting projections of multi-dimensional spaces.

D) None of these

Answer: A

46) Query tools are

A) A reference to the speed of an algorithm, which is quadratically dependent on the size of the data

B) Attributes of a database table that can take only numerical values.

C) Tools designed to query a database.

D) None of these

Answer: C

47) Operational database is

A) A measure of the desired maximal complexity of data mining algorithms

B) A database containing volatile data used for the daily operation of an organization

C) Relational database management system

D) None of these

Answer: B

48) Which of the following is/are the Data mining tasks?

(a) Regression

(b) Classification

(c) Clustering

(d) inference of associative rules

(e) All (a), (b), (c) and (d) above.

Answer: E

Explanation: Regression, Classification and Clustering are the data mining tasks.

49) In a data warehouse, if D1 and D2 are two conformed dimensions, then

(a) D1 may be an exact replica of D2

(b) D1 may be at a rolled up level of granularity compared to D2

(c) Columns of D1 may be a subset of D2 and vice versa

(d) Rows of D1 may be a subset of D2 and vice versa

(e) All (a), (b), (c) and (d) above.

Answer: A

Explanation: In a data warehouse, if D1 and D2 are two conformed dimensions, then D1 may be an exact replica of D2.

50. Which of the following is not an ETL tool?

(a) Informatica

(b) Oracle warehouse builder

(c) Datastage

(d) Visual studio

(e) DT/studio.

Answer: D

Explanation: Visual Studio is not an ETL tool.

51) .... 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

Answer: B) Data mining

52) 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

Answer: D) All i, ii, iii, iv and v

53) 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

Answer: C) Selection and interpretation

54) .... 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

Answer: A) Data Characterization

55) .... 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

Answer: C) Data discrimination

56) Strategic value of data mining is ....

A) cost-sensitive

B) work-sensitive

C) time-sensitive

D) technical-sensitive

Answer: C) time-sensitive

57) .... 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

Answer: B) Data Classification

58. 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

Answer: D) All i, ii, iii and iv

59) The full form of KDD is ....

A) Knowledge Database

B) Knowledge Discovery Database

C) Knowledge Data House

D) Knowledge Data Definition

Answer: B) Knowledge Discovery Database

60) The out put of KDD is ....

A) Data

B) Information

C) Query

D) Useful information

Answer: D) Useful information

61. The full form of OLAP is

A) Online Analytical Processing

B) Online Advanced Processing

C) Online Advanced Preparation

D) Online Analytical Performance

Answer: A) Online Analytical Processing

62) .... 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

Answer: B) Data Warehousing

63) The data is stored, retrieved and updated in ....

A) OLAP

B) OLTP

C) SMTP

D) FTP

Answer: B) OLTP

64) 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

Answer: A) OLAP

65) .... is a good alternative to the star schema.

A) Star schema

B) Snowflake schema

C) Fact constellation

D) Star-snowflake schema

Answer: C) Fact constellation

66) 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

Answer: C) data source view

67) The type of relationship in star schema is ....

A) many to many

B) one to one

C) one to many

D) many to one

Answer: C) one to many

68) 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

Answer: A) top-down view

69) 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

Answer: D) Component Key

70) Which of the following is not a kind of data warehouse application?

A) Information processing

B) Analytical processing

C) Data mining

D) Transaction processing

Answer: D) Transaction processing

71) Data warehouse architecture is based on ....

A) DBMS

B) RDBMS

C) Sybase

D) SQL Server

Answer:B) RDBMS

72) .... 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

Answer: B) Analytical processing

73) 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

Answer: C) Data cube

74) 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

Answer: A) Current detail data

75) 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

Answer: C) organized around important subject areas

76) 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

Answer: D) All of the above

77) Data warehouse contains .... data that is never found in the operational environment.

A) normalized

B) informational

C) summary

D) denormalized

Answer: C) summary

78) .... are responsible for running queries and reports against data warehouse tables.

A) Hardware

B) Software

C) End users

D) Middle ware

Answer: C) End users

79) The biggest drawback of the level indicator in the classic star schema is that is limits ....

A) flexibility

B) quantify

C) qualify

D) ability

Answer: A) flexibility

80) .... 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

Answer: C) Multidimensional database

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

(a) XML data exchange standards

(b) Data access tools to be used

(c) Metadata naming conventions

(d) Extract, Transform, and Load (ETL) tool to be used

(e) All (a), (b), (c) and (d) above.

Answer: B

Explanation: Data access tools to be used when deciding on the data structure of a data mart.

82) 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

(e) Compression of data.

Answer: C

Explanation: The process of removing the deficiencies and loopholes in the data is called as cleaning up of data.

83) Which one manages both current and historic transactions?

(a) OLTP

(b) OLAP

(c) Spread sheet

(d) XML

(e) All (a), (b), (c) and (d) above.

Answer: B

Explanation: Online Analytical Processing (OLAP) manages both current and historic transactions.

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

(a) Partitioning

(b) Grid

(c) Cluster

(d) Table

(e) Data source.

Answer: C

Explanation: Cluster is the collection of data objects that are similar to one another within the same group.

85) Which of the following employees data mining techniques to analyze the intent of a user query, provided additional generalized or associated information relevant to the query?

(a) Iceberg query method

(b) Data analyzer

(c) Intelligent query answering

(d) DBA

(e) Query parser.

Answer: C

Explanation: Intelligent Query Answering   employee's data mining techniques to analyze the intent of a user query provided additional generalized or associated information relevant to the query.

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

(a) KDD process

(b) ETL process

(c) KTL process

(d) MDX process

(e) None of the above.

Answer: A

Explanation: KDD Process includes data cleaning, data integration, data selection, data transformation, data mining, pattern evolution, and knowledge presentation.

87. At which level we can create dimensional models?

(a) Business requirements level

(b) Architecture models level

(c) Detailed models level

(d) Implementation level

(e) Testing level.

Answer: B

Explanation: Dimensional models can be created at Architecture models level.

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

(a) Verbose

(b) Descriptive

(c) Equally unavailable

(d) Complete

(e) Indexed.

Answer: C

Explanation: Equally unavailable is not related to dimension table attributes.

89) Data warehouse bus matrix is a combination of

(a) Dimensions and data marts

(b) Dimensions and facts

(c) Facts and data marts

(d) Dimensions and detailed facts

(e) All (a), (b), (c) and (d) above.

Answer: A

Explanation: Data warehouse bus matrix is a combination of Dimensions and data marts.

90) Which of the following is not the managing issue in the modeling process?

(a) Content of primary units column

(b) Document each candidate data source

(c) Do regions report to zones

(d) Walk through business scenarios

(e) Ensure that the transaction edit flat is used for analysis.

Answer: E

Explanation: Ensure that the transaction edit flat is used for analysis is not the managing issue in the modeling process.

91) Data modeling technique used for data marts is

(a) Dimensional modeling

(b) ER – model

(c) Extended ER – model

(d) Physical model

(e) Logical model.

Answer: A

Explanation: Data modeling technique used for data marts is Dimensional modeling.

92) A warehouse architect is trying to determine what data must be included in the warehouse. A meeting has been arranged with a business analyst to understand the data requirements, which of the following should be included in the agenda?

(a) Number of users

(b) Corporate objectives

(c) Database design

(d) Routine reporting

(e) Budget.

Answer: D

Explanation: Routine reporting should be included in the agenda.

93. An OLAP tool provides for

(a) Multidimensional analysis

(b) Roll-up and drill-down

(c) Slicing and dicing

(d) Rotation

(e) Setting up only relations.

Answer: C

Explanation: An OLAP tool provides for Slicing and dicing.

94. The Synonym for data mining is

(a) Data warehouse

(b) Knowledge discovery in database

(c) ETL

(d) Business intelligence

(e) OLAP.

Answer: C

Explanation: The synonym for data mining is Knowledge discovery in Database.

95) Which of the following statements is true?

(a) A fact table describes the transactions stored in a DWH

(b) A fact table describes the granularity of data held in a DWH

(c) The fact table of a data warehouse is the main store of descriptions of the transactions stored in a DWH

(d) The fact table of a data warehouse is the main store of all of the recorded transactions over time

(e) A fact table maintains the old records of the database.

Answer: D

Explanation: The fact table of a data warehouse is the main store of all of the recorded transactions over time is the correct statement.

96) Most common kind of queries in a data warehouse

(a) Inside-out queries

(b) Outside-in queries

(c) Browse queries

(d) Range queries

(e) All (a), (b), (c) and (d) above.

Answer: A

Explanation: The Most common kind of queries in a data warehouse is Inside-out queries.

97) Concept description is the basic form of the

(a) Predictive data mining

(b) Descriptive data mining

(c) Data warehouse

(d) Relational data base

(e) Proactive data mining.

Answer: B

Explanation: Concept description is the basis form of the descriptive data mining.

98) The apriori property means

(a) If a set cannot pass a test, all of its supersets will fail the same test as well

(b) To improve the efficiency the level-wise generation of frequent item sets

(c) If a set can pass a test, all of its supersets will fail the same test as well

(d) To decrease the efficiency the level-wise generation of frequent item sets

(e) All (a), (b), (c) and (d) above.

Answer: B

Explanation: The apriori property means to improve the efficiency the level-wise generation of frequent item sets.

99) Which of following form the set of data created to support a specific short lived business situation?

(a) Personal data marts

(b) Application models

(c) Downstream systems

(d) Disposable data marts

(e) Data mining models.

Answer: D

Explanation: Disposable Data Marts is the form the set of data created to support a specific short lived business situation.

100) What is/are the different types of Meta data?

I. Administrative.

II. Business.

III. Operational.

(a) Only (I) above

(b) Both (II) and (III) above

(c) Both (I) and (II) above

(d) Both (I) and (III) above

(e) All (I), (II) and (III) above.

Answer: E

Explanation: The different types of Meta data are Administrative, Business and Operational.

101) Multiple Regression means

(a) Data are modeled using a straight line

(b) Data are modeled using a curve line

(c) Extension of linear regression involving only one predicator value

(d) Extension of linear regression involving more than one predicator value

(e) All (a), (b), (c) and (d) above.

Answer: D

Explanation: Multiple Regression means extension of linear regression involving more than one predicator value.

102) Which of the following should not be considered for each dimension attribute?

(a) Attribute name

(b) Rapid changing dimension policy

(c) Attribute definition

(d) Sample data

(e) Cardinality.

Answer: B

Explanation: Rapid changing dimension policy should not be considered for each dimension attribute.

103) A Business Intelligence system requires data from:

(a) Data warehouse

(b) Operational systems

(c) All possible sources within the organization and possibly from external sources

(d) Web servers

(e) Database servers.

Answer: A

Explanation: A business Intelligence system requires data from Data warehouse

104) Data mining application domains are

(a) Biomedical

(b) DNA data analysis

(c) Financial data analysis

(d) Retail industry and telecommunication industry

(e) All (a), (b), (c) and (d) above.

Answer: E

Explanation: Data mining application domains are Biomedical, DNA data analysis, Financial data analysis and Retail industry and telecommunication industry

105. The generalization of multidimensional attributes of a complex object class can be performed by examining each attribute, generalizing each attribute to simple-value data and constructing a multidimensional data cube is called as

(a) Object cube

(b) Relational cube

(c) Transactional cube

(d) Tuple

(e) Attribute.

Answer: A

Explanation: The generalization of multidimensional attributes of a complex object class can be performed by examining each attribute, generalizing each attribute to simple-value data and constructing a multidimensional data cube is called as object cube.

106. Which of the following project is a building a data mart for a business process/department that is very critical for your organization?

(a) High risk high reward

(b) High risk low reward

(c) Low risk low reward

(d) Low risk high reward

(e) Involves high risks.

Answer: A

Explanation: High risk high reward project is a building a data mart for a business process/department that is very critical for your organization

107. Which of the following tools a business intelligence system will have?

(a) OLAP tool

(b) Data mining tool

(c) Reporting tool

(d) Both(a) and (b) above

(e) (a), (b) and (c) above.

Answer: A

Explanation: Business intelligence system will have OLAP, Data mining and reporting tolls.