

PRIST DEEMED UNIVERSITY ONLINE EXAMINATIONS - 2020

10.09.2020 AN_16151C73 - DATA WARE HOUSING & DATA MINING

INFORMATION TECHNOLOGY - IV YEAR _ VII SEMESTER

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 the mentioned

Clear selection

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 the mentioned

Clear selection



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.

Clear selection

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

Clear selection

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

Clear selection



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

Clear selection

Heterogeneous databases 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 the mentioned

Clear selection

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

Clear selection



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

- ☐ a) subject-oriented
- ☐ b) time-variant
- ☐ c) integrated.
- ☒ d) All of the above

Clear selection

The extract process is which of the following?

- ☐ a) Capturing all of the data contained in various operational systems
- ☒ b) Capturing a subset of the data contained in various operational systems
- ☐ c) Capturing all of the data contained in various decision support systems
- ☐ d) Capturing a subset of the data contained in various decision support systems

Clear selection

_____ 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.

Clear selection



Data dictionary is

- ☐ a) Large collection of data mostly stored in a computer system
- ☐ b) The removal of noise errors and incorrect input from a database
- ☒ c) The systematic description of the syntactic structure of a specific database. It describes the structure of the attributes the tables and foreign key relations hips.
- ☐ d) None of these

Clear selection

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 the mentioned

Clear selection



Transient data is which of the following?

- ☒ a) Data in which changes to existing records cause the previous version of the records to be eliminated
- ☐ b) Data in which changes to existing records do not cause the previous version of the records to be eliminated
- ☐ c) Data that are never altered or deleted once they have been added
- ☐ d) Data that are never deleted once they have been added

Clear selection

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

Clear selection

Data cleaning is

- ☐ a) Large collection of data mostly stored in a computer system
- ☒ b) The removal of noise errors and incorrect input from a database
- ☐ d) None of these
- ☐ c) The systematic description of the syntactic structure of a specific database. It describes the structure of the attributes the tables and foreign key relationships.

Clear selection



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

- ☐ a) Fact
- ☒ d) All of the above
- ☐ c) Helper
- ☐ b) Dimension

Clear selection

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 the mentioned

Clear selection

The data is stored, retrieved & updated in _____

- ☐ a) OLAP
- ☒ b) OLTP.
- ☐ c) SMTP
- ☐ d) FTP

Clear selection



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

Clear selection

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 mentioned

Clear selection

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 the mentioned

Clear selection



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

Clear selection

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) All of the mentioned

Clear selection

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

Clear selection



Fact tables are which of the following?

- ☐ a) Completely denormalized
- ☐ b) Partially denormalized
- ☒ c) Completely normalized
- ☐ d) Partially normalized

Clear selection

The data Warehouse is _____

- ☒ a) read only.
- ☐ c) read write only.
- ☐ b) write only.
- ☐ d) none.

Clear selection

Reconciled data is which of the following?

- ☐ a) Data stored in the various operational systems throughout the organization.
- ☒ b) Current data intended to be the single source for all decision support systems.
- ☐ c) Data stored in one operational system in the organization.
- ☐ d) Data that has been selected and formatted for end-user support applications.

Clear selection



A multifield transformation does which of the following?

- ☐ a) Converts data from one field into multiple fields
- ☐ b) Converts data from multiple fields into one field
- ☐ c) Converts data from multiple fields into multiple fields
- ☒ d) All of the above

Clear selection

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

Clear selection

Decision support systems (DSS) is

- ☐ a) A family of relational database management systems marketed by IBM
- ☒ b) Interactive systems that enable decision makers to use databases and models on a computer in order to solve ill-structured problems
- ☐ d) None of these
- ☐ c) It consists of nodes and branches starting from a single root node. Each node represents a test, or decision

Clear selection



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

Clear selection

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 the mentioned

Clear selection

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

Clear selection



The time horizon in Data warehouse is usually _____.

- ☐ a) 1-2years
- ☒ d) 5-10years
- ☐ c) 5-6years
- ☐ b) 3-4years

Clear selection

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

Clear selection

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

Clear selection



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

Clear selection

The load and index 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

Clear selection

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

Clear selection



Back

Submit

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#) - [Privacy Policy](#).

Google Forms

