

1. In the decision making process proposed by Simon (1977), which of the following should be the second phase:
- A) implementation
B) Intelligence
C) Design
D) Choice
- Int.
Design
Choice
+ implementation.
2. A system that relies mostly on heuristics (what-if-then):
- A) Enterprise Resource Planning
B) Expert System
C) Executive Information Systems
D) Data Warehousing
3. A system that relies mostly on historical and comprehensive data:
- A) Enterprise Resource Planning
B) Expert System
C) Executive Information Systems
D) Data Warehousing
4. A system that relies mostly on providing dashboards and easy to understand graphics:
- A) Enterprise Resource Planning ~
B) Expert System
C) Executive Information Systems ~
D) Data Warehousing x
5. A system that creates a single-source of truth for data items on daily functional activities:
- A) Enterprise Resource Planning
B) Expert System
C) Executive Information Systems
D) Data Warehousing
- is Redan.

1. One of the most important systems that make business intelligence possible is:
- A) Enterprise Resource Planning.
 - B) Expert System.
 - C) Executive Information Systems.
 - ☒ D) Data Warehousing.
2. What is a data mart:
- E) A web-based database where clients can see their transaction history.
 - F) A system where companies are able to buy data about customers.
 - ☒ G) A smaller sub-set of a data warehouse dedicated for a special business function.
 - H) A much larger system than a data warehouse that includes external data.
3. The letters in the ETL process stand for:
- ☒ A) Extract, Transform, Load.
 - B) Enhance, Toggle, Load.
 - C) Extend, Track, Load.
 - D) Extract, Terminate, Load.
4. The purpose of descriptive analytics is to:
- A) Give recommendations on what actions to take.
 - B) Determine the expected future outcome of events.
 - ☒ C) Describe the current situation.
 - D) Take automated actions driven by AI.
5. The purpose of predictive analytics is to:
- A) Give recommendations on what actions to take.
 - ☒ B) Determine the expected future outcome of events.
 - C) Describe the current situation.
 - D) Take automated actions driven by AI.

1. Which of the following is NOT a biometric related AI technology:
A) Natural language processing.
B) Industrial robotics.
C) Image recognition.
D) Voice recognition. ✓
2. Which of the following is NOT an AI application:
E) Smart homes.
F) Prediction.
G) Robotics.
H) Point-of-sales transaction processing. ✓
3. Which of the following is an example of strong AI applications:
A) Chess game.
B) Medical diagnosis.
C) Autonomous vehicles.
D) Automated call centers. ✓
4. Monitoring systems are an example of which of the following:
A) Assisted intelligence.
B) Autonomous AI.
C) Augmented intelligence.
D) Strong AI. ✓
5. Which of the following is not a main areas/theories used in Analytics:
A) Statistics. ✓
B) Cognitive computing. ✓
C) Operations research. ✓
D) Management science. ✓

1. Which of the following areas can be considered part of data mining:

- A) Management Science.
- B) Artificial Intelligence. ✓
- C) Information visualization. ✓
- D) All options are true.

2. What was/were the event(s) that helped the Miami Dade police department accept data mining in their crime detection operations:

- A) A much higher crime rate than the last year.
- B) The recommendation by the state government to start using technology in crime fighting.
- ~~C) The solving of some previously unsolved crimes.~~
- D) All options are true.

3. Which of the following is/are important source(s) of data for data mining:

- A) A data warehouse. ✓
- B) Web log files.
- C) System data bases. ✓
- D) All options are true.

4. In association types of patterns:

- A) Future expected events or values are determined.
- B) The natural grouping of things based on their characteristics.
- C) The grouping of related things.
- D) The determination of things to come in order.

5. In segmentation types of patterns:

- A) Future expected events or values are determined.
- B) The natural grouping of things based on their characteristics.
- C) The grouping of related things.
- D) The determination of things to come in order.

1. According to the taxonomy of data mining as defined in the textbook, which of the following is NOT a data mining task:
 - A) Prediction.
 - B) Segmentation.
 - C) Association.
 - D) Link analysis.
2. According to the taxonomy of data mining as defined in the textbook, which of the following is NOT a data mining method:
 - A) Regression.
 - B) Knowledge discovery.
 - C) Clustering.
 - D) Market-basket.
3. When it comes to types of learning in a data mining model, unsupervised learning involves which of the following:
 - A) Providing only output values to the model.
 - B) Providing both input and output values to the model.
 - C) Providing only input values to the model.
 - D) Noe of the options is true.
4. Which data mining method is used to determine the future expected value of the same variable:
 - A) Link analysis.
 - B) Market-basket.
 - C) Sequence analysis.
 - D) Time series.
5. Which data mining method is used to discover relationships between groups of people or things:
 - A) Link analysis.
 - B) Market-basket.
 - C) Sequence analysis.
 - D) Time series.

1. According to the taxonomy of data mining as defined in the textbook, which of the following is NOT a data mining task:
A) Prediction.
B) Segmentation.
C) Association.
D) Link analysis.
2. According to the taxonomy of data mining as defined in the textbook, which of the following is NOT a data mining method:
A) Regression.
B) Knowledge discovery.
C) Clustering.
D) Market-basket.
3. When it comes to types of learning in a data mining model, unsupervised learning involves which of the following:
A) Providing only output values to the model.
B) Providing both input and output values to the model.
C) Providing only input values to the model.
D) None of the options is true.
4. Which data mining method is used to determine the future expected value of the same variable:
A) Link analysis.
B) Market-basket.
C) ~~Sequence analysis.~~
D) Time series.
5. Which data mining method is used to discover relationships between groups of people or things:
A) Link analysis.
B) Market-basket.
C) Sequence analysis.
D) Time series.

1. Which of the data mining processes involves a step that tries to get a good idea about the objective of the desired data mining model to be developed:

- A) Classification.
- B) CRISP-DM.
- C) SEMMA.
- D) KDD.

2. Which of the data mining processes is considered to be a more comprehensive process which actually involves a data-mining step:

- A) Classification.
- B) CRISP-DM.
- C) SEMMA.
- D) KDD.

3. In the CRISP-DM data mining process the phase that takes the most of the project time is:

- A) Data Understanding
- B) Data Preparation
- C) Model Building
- D) Testing and Evaluation

4. The most used data mining method is:

- A) Clustering.
- B) Classification.
- C) Time-series.
- D) Regression.

5. Which of the following is NOT an example of nominal data:

- A) High, middle, low
- B) Open, closed.
- C) Blue, green, yellow.
- D) Pen, pencil, marker.

1. Which of the following estimation methodology for classification produces more accurate prediction?
☒ X
A) Single split.
B) K-fold cross-validation.
C) Confusion matrix
D) All options are true.
2. What is a different name for single split estimation methodology?
☒ ✓
A) Simple split.
B) Holdout.
C) Sample test estimation
D) All options are true.
3. In single test estimation methodology, what percent of the data is used for training the model:
☒ ✓
A) 1/2.
B) 1/3.
C) 1/4.
D) 2/3.
4. If $k=5$ in a K-fold cross-validation methodology, what percent of the data is used as testing data in each run:
☒ ✓
A) 10%.
B) 80%.
C) 20%.
D) 5%.
 $\frac{5}{5}$
5. What is the main purpose of using a confusion matrix?
☒ ✓
A) Determine which data to use for training and testing the model
B) Determine the accuracy of the developed classification model
C) Determine if there is error or missing data
D) All options are true