

Question #1 of 9

Question ID: 1572907

Under which of these conditions is a machine learning model said to be underfit?

- A) The model treats true parameters as noise. 
- B) The input data are not labeled. 
- C) The model identifies spurious relationships. 

Explanation

Underfitting describes a machine learning model that is not complex enough to describe the data it is meant to analyze. An underfit model treats true parameters as noise and fails to identify the actual patterns and relationships. A model that is overfit (too complex) will tend to identify spurious relationships in the data. Labeling of input data is related to the use of supervised or unsupervised machine learning techniques.

(Module 11.1, LOS 11.b)

Question #2 of 9

Question ID: 1572904

An executive describes her company's "low latency, multiple terabyte" requirements for managing Big Data. To which characteristics of Big Data is the executive referring?

- A) Velocity and variety. 
- B) Volume and variety. 
- C) Volume and velocity. 

Explanation

Big Data may be characterized by its volume (the amount of data available), velocity (the speed at which data are communicated), and variety (degrees of structure in which data exist). "Terabyte" is a measure of volume. "Latency" refers to velocity.

(Module 11.1, LOS 11.b)

Question #3 of 9

Question ID: 1572909

A data analyst uses fintech to evaluate the number of times the words *buy* or *sell* appear in a company's quarterly filings in a given fiscal year. This is *most likely* an example of which form of fintech?

- A) Text analytics. 
- B) Natural language processing. 
- C) Algorithmic trading. 

Explanation

Text analytics, which relates to the analysis of unstructured data in text or voice forms, can be used to analyze the frequency in which a word or words appear in documents. The number of times the words *buy* and *sell* appear can be evaluated using text analytics. Algorithmic trading is computerized securities trading based on preset trading rules. Natural language processing uses computers and artificial intelligence (AI) to interpret human language.

(Module 11.1, LOS 11.c)

Question #4 of 9

Question ID: 1572903

Which of the following statements about fintech is *most accurate*?

- A) A primary driver of fintech is the increasingly structured nature of data that firms must process. 
- B) Financial services that involve subjective judgment, such as investment advice, are unlikely to be affected by fintech. 
- C) Fintech companies include those that develop technology for the financial services industry. 

Explanation

Fintech refers to technological developments with potential applications in financial services, as well as to the industry that develops these technologies. While firms must process an increasing volume of data, a large portion of that data exists in unstructured forms. Automated investment advice is a potential application of fintech.

(Module 11.1, LOS 11.a)

Question #5 of 9

Question ID: 1572911

Which of the following statements *most accurately* describes a data processing method?

- A) Capture focuses on how data moves from the underlying source to the analytical tool. 
- B) Curation focuses on data quality and accuracy through data cleaning. 
- C) Search focuses on how data will be recorded and archived. 

Explanation

Curation refers to ensuring the quality and accuracy of data. *Capture* refers to collecting and transforming data in preparation for analysis. *Search* refers to the ways data will be queried.

(Module 11.1, LOS 11.c)

Question #6 of 9

Question ID: 1572910

A large investment company uses an enterprise risk management framework to assess the various risks in its organization. Some of the tools it uses to assess its risks include scenario analysis and simulations, which typically involve:

- A) small amounts of quantitative and qualitative data. 
- B) large amounts of quantitative data and small amounts of qualitative data. 
- C) large amounts of quantitative and qualitative data. 

Explanation

The techniques (e.g., simulations and scenario analysis) used to assess and manage risk will require large amounts of quantitative and qualitative data. This is particularly true for a large investment company.

(Module 11.1, LOS 11.c)

Question #7 of 9

Question ID: 1572905

Which of the following uses of data is *most accurately* described as curation?

- A) An analyst adjusts daily stock index data from two countries for their different market holidays. 
- B) A data technician accesses an offsite archive to retrieve data that has been stored there. 

- C) An investor creates a word cloud from financial analysts' recent research reports about a company. 

Explanation

Curation is ensuring the quality of data—for example, by adjusting for bad or missing data. Word clouds are a visualization technique. Moving data from a storage medium to where they are needed is referred to as transfer.

(Module 11.1, LOS 11.b)

Question #8 of 9

Question ID: 1572908

The technique in which a machine learns to model a set of output data from a given set of inputs is *best* described as:

- A) deep learning. 
- B) supervised learning. 
- C) unsupervised learning. 

Explanation

Supervised learning is a machine learning technique in which a machine is given labeled input and output data and then models the output data based on the input data. In unsupervised learning, a machine is given input data in which to identify patterns and relationships, but no output data to model. Deep learning is a technique to identify patterns of increasing complexity, and may use supervised or unsupervised learning.

(Module 11.1, LOS 11.b)

Question #9 of 9

Question ID: 1572906

Artificial intelligence is *best* described as:

- A) networks of smart devices and buildings. 
- B) the field of study concerned with extracting information from data. 
- C) computer systems that emulate human thinking. 

Explanation

Artificial intelligence refers to computer systems that emulate the functioning of the human mind. Networks of smart devices and buildings are referred to as the Internet of Things. Data science is the field of study concerned with extracting information from data.

(Module 11.1, LOS 11.b)