

11

INTRODUCTION TO BIG
DATA

- Under which of these conditions is a machine learning model said to be underfit?
 - The model treats true parameters as noise.
 - The input data are not labelled.
 - The model identifies spurious relationships.
- 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?
 - Velocity and variety.
 - Volume and variety.
 - Volume and velocity.
- 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?
 - Text analytics.
 - Natural language processing.
 - Algorithmic trading
- Which of the following statements about fintech is most accurate?
 - A primary driver of fintech is the increasingly structured nature of data that firms must process.
 - Financial services that involve subjective judgment, such as investment advice, are unlikely to be affected by fintech.
 - Fintech companies include those that develop technology for the financial services industry.
- Which of the following statements most accurately describes a data processing method?
 - Capture focuses on how data moves from the underlying source to the analytical tool.
 - Curation focuses on data quality and accuracy through data cleaning.
 - Search focuses on how data will be recorded and archived.

6. 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.
7. 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.
8. 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.
9. 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.

