

# STUDY SESSION

# 3

## Quantitative Methods (2)

**T**his study session provides coverage on techniques that underlie how financial technology (fintech) is affecting areas within the investment industry, such as investment analysis, automated advice, and risk management. The first reading introduces techniques in machine learning (ML) that involve clustering, simplifying, classifying, and predicting relationships in the large datasets that are often found in finance. The next reading examines how data projects involving large datasets are structured with an application to sentiment analysis in investment analysis using machine learning techniques for natural language processing (NLP). The session concludes with coverage of probability-based techniques for assessing risk, with a focus on simulation models.

### READING ASSIGNMENTS

|                  |  |
|------------------|--|
| <b>Reading 7</b> | Machine Learning<br>by Kathleen DeRose, CFA, and Christophe Le Lannou  |
| <b>Reading 8</b> | Big Data Projects<br>by Sree Mallikarjun, PhD, and Ahmed Abbasi, PhD   |
| <b>Reading 9</b> | Excerpt from “Probabilistic Approaches: Scenario Analysis, Decision Trees, and Simulations”<br>by Aswath Damodaran |

### LEARNING OUTCOMES

#### READING 7. MACHINE LEARNING

The candidate should be able to:

- a** distinguish between supervised machine learning, unsupervised machine learning, and deep learning;

- b** describe overfitting and identify methods of addressing it;
- c** describe supervised machine learning algorithms—including penalized regression, support vector machine, k-nearest neighbor, classification and regression tree, ensemble learning, and random forest—and determine the problems for which they are best suited;
- d** describe unsupervised machine learning algorithms—including principal components analysis, k-means clustering, and hierarchical clustering—and determine the problems for which they are best suited;
- e** describe neural networks, deep learning nets, and reinforcement learning.

## READING 8. BIG DATA PROJECTS

The candidate should be able to:

- a** state and explain steps in a data analysis project;
- b** describe objectives, steps, and examples of preparing and wrangling data;
- c** describe objectives, methods, and examples of data exploration;
- d** describe objectives, steps, and techniques in model training;
- e** describe preparing, wrangling, and exploring text-based data for financial forecasting;
- f** describe methods for extracting, selecting and engineering features from textual data;
- g** evaluate the fit of a machine learning algorithm.

## READING 9. EXCERPT FROM “PROBABILISTIC APPROACHES: SCENARIO ANALYSIS, DECISION TREES, AND SIMULATIONS”

The candidate should be able to:

- a** describe steps in running a simulation;
- b** explain three ways to define the probability distributions for a simulation’s variables;
- c** describe how to treat correlation across variables in a simulation;
- d** describe advantages of using simulations in decision making;
- e** describe some common constraints introduced into simulations;
- f** describe issues in using simulations in risk assessment;
- g** compare scenario analysis, decision trees, and simulations.