DLI Accelerated Data Science Teaching Kit

Module 18 Quiz:
Streaming Data

1. Compared to general data, data preprocessing is much easier for streaming data.

A. True
B. False

Answer: B

Explanation: See lecture 18.2 slides

2. A 'sketch' can avoid storing and maintaining a large amount of data.

A. True
B. False

Answer: A

Explanation: See lecture 18.2 slides

3. In contrast to dimensionality reduction, feature selection does not apply transformations to the data, and thus, the features can still be interpreted.

A. True
B. False

Answer: A

Explanation: See lecture 18.2 slides

4. Which challenge will be involved during learning on streaming data?

A. Building the relationship between data streams and time series
B. Addressing the problem of dealing with partially and delayed labels
C. Learning on imbalanced data streams
D. Detecting anomalies from streaming data

Answer: ABCD

Explanation: See lecture 18.3 slides

5. Semi-supervised learning (SSL) is particularly relevant to streaming applications where data are abundant but labeled data may be rare.

A. True
B. False

Answer: A

Explanation: See lecture 18.3 slides

6. Learning systems on streaming data should be able to monitor their working conditions since they act in dynamic environments, where working conditions change and evolve.

A. True
B. False

Answer: A

Explanation: See lecture 18.4 slides

7. Which change is related to feature drift?

A. Change in the values of a feature and their association with the class
B. Change in the domain of features
C. Change in the subset of features that are used to label an instance
D. Change in learning system

Answer: ABC

Explanation: See lecture 18.4 slides

8. Hyperparameter tuning for streaming data analysis is time-consuming and labor-intensive.

A. True
B. False

Answer: A

Explanation: See lecture 18.4 slides