

PRACTICE QUESTIONS

Snowflake

SnowPro Advanced: MLOps Engineer

SHARED BY [u/bright_wombat_29](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 15, 2026

EXAM Snowflake SnowPro Advanced: MLOps Engineer

Shared by [u/bright_wombat_29](#) on aiexamcerts

10 practice questions for the Snowflake SnowPro Advanced: MLOps Engineer exam, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/snowflake/snowpro-advanced-mlops-engineer. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. Which compute option is most appropriate for memory-heavy, in-warehouse model training on a large dataset?
 - A A reader account that is used only for dashboards and reports
 - B Snowpark-optimized warehouses, which provide extra memory for Python workloads
 - C Multi-cluster standard warehouses set to the smallest size, with a long auto-suspend delay to keep them warm
 - D X-Small standard warehouses with auto-suspend set to one minute

Answer: B

Snowpark-optimized warehouses are built with more memory per node, which suits training code that holds large objects in memory. A standard warehouse of any size, including a multi-cluster one, offers the same memory profile per node, which makes the smallest size the tempting wrong choice. A reader account is for sharing data and does not run training.

2. A feature table keyed by entity and event time, queried for the latest values as of a given moment, is accessed with a _____ join.

Answer: point-in-time (also accepted: point in time, as-of)

A point-in-time or as-of join picks, for each row, the most recent feature value at or before its timestamp. An ordinary join on the key alone can pick values from after the event.

3. Which Snowflake object defines the set of nodes that Snowpark Container Services workloads run on?
 - A A stream that records row changes on a source table
 - B A masking policy that hides column values from some roles
 - C A compute pool with a chosen instance family and minimum and maximum node counts
 - D A resource monitor that caps the credits used by a warehouse

Answer: C

A compute pool defines the nodes, instance family and node limits used by container services, which makes it the right object for hosting serving or training containers. A resource monitor controls credit spending, a stream records changes, and a masking policy hides values, so each is a real Snowflake object that answers a different question.

4. Imputing missing values with statistics computed on the full dataset, including the test split, is a common source of data leakage.

- A True
- B False

Answer: A (True)

Using the test rows to compute the imputation value lets information about the held-out data shape the training features, which inflates evaluation results. Imputation should be fit on the training fold only and applied to the other folds. The statement is true.

5. Which two details should be recorded for each feature to support reuse? (Choose two.)

- A The IP address of the analyst who first created it in the account
- B The source tables, transformations and refresh frequency
- C The number of rows in the latest training table built from it
- D The owner and the business meaning of the feature
- E The cluster key of the raw source table it was derived from

Answer: B, D

Owner and meaning let others find and trust a feature, while sources, transformations and refresh frequency let them judge its freshness and reproduce it. The other options record details that do not help anyone decide whether to reuse a feature.

6. A categorical feature contains a value that never appeared in training. What should the production pipeline do?

- A Map the unseen value to a defined fallback category that was agreed during training
- B Retrain the model inside the inference query for each request
- C Stop the pipeline and drop every row in the source table that contains an unseen value at all
- D Create a new feature column at run time for every unseen value

Answer: A

A defined fallback category, such as an unknown bucket, keeps inference running predictably and makes the handling of new values explicit. Dropping all rows would halt scoring, retraining per request is impractical, and creating columns at run time breaks the feature schema the model expects. Each of those fails in a way that is harder to diagnose.

7. A model reaches 99 percent training accuracy but only 70 percent validation accuracy. What is the most likely issue?
- A Underfitting caused by having too many training rows available
 - B Overfitting to the training data
 - C A missing primary key on the validation table in the database
 - D A data residency violation in the training region chosen

Answer: B

A large gap between training and validation performance is the classic sign of overfitting, where the model memorizes training patterns that do not generalize. Underfitting shows poor scores on both sets, so having too many rows is not its cause. Residency and keys do not explain a performance gap.

8. Searching many hyperparameter combinations automatically to find the best settings is called hyperparameter _____.

Answer: tuning (also accepted: optimization)

Hyperparameter tuning evaluates candidate settings, often with a search strategy, and keeps the best by validation performance. The search should use validation data, not the final test set.

9. Which option serves a model for real-time REST calls within the Snowflake ecosystem?
- A A Streamlit app that displays the model's training loss over time
 - B A listing offered to external accounts for read-only data access
 - C A scheduled task that writes a CSV file of predictions every night
 - D Model serving on Snowpark Container Services

Answer: D

Model serving on Snowpark Container Services hosts the model in a container that exposes an endpoint for real-time calls. A nightly task supports batch output, a listing shares data and a Streamlit dashboard visualizes metrics, so none of them serves live requests on its own.

10. Which deployment risk does a canary release specifically reduce?
- A Exposure of all users to a faulty model version at the same moment
 - B Long compile times for the feature SQL in the development environment
 - C Storage cost of keeping the training data for an extra month
 - D The need for any monitoring of the model after it goes live

Answer: A

A canary sends a small share of traffic to the new version first, so a fault affects few users and can be caught before a full switch. It does not change storage, monitoring needs or compile times, which are separate concerns.