

PRACTICE QUESTIONS

Microsoft

AI-300

Machine Learning Operations Engineer Associate

SHARED BY [u/Mellow-Ferret-8106](#) on aiexamcerts

EDITION	Free Set 2
QUESTIONS	10
SHARED	October 15, 2026
EXAM	Microsoft AI-300, Machine Learning Operations Engineer Associate

Shared by u/Mellow-Ferret-8106 on aiexamcerts

10 practice questions for the Microsoft AI-300 exam, Machine Learning Operations Engineer Associate, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/microsoft/ai-300. Spotted a mistake? Email support@aiexamcerts.com with the exam code AI-300 and the question number.

1. A team needs to share registered models and environments with a workspace in another region. Which feature supports sharing assets across workspaces?

- A Local compute instance snapshots
- B Azure Machine Learning registry
- C Datastore SAS tokens
- D Workspace Git sync

Answer: B

An Azure Machine Learning registry shares models, environments, and components across workspaces and regions. Datastore SAS tokens are the tempting shortcut, but they grant file access and skip versioning and the registered metadata. Snapshots and Git sync do not share assets between workspaces.

2. Deleting a data asset version in Azure Machine Learning also deletes the underlying files in the datastore.

- A True
- B False

Answer: B (False)

Deleting an asset version removes the registration, while the underlying storage files remain. Assuming that the registration owns the files is the tempting mistake, and it can leave orphaned data in storage.

3. Which setting bounds the total number of trials that a hyperparameter sweep job runs?

- A The number of endpoint replicas deployed
- B The compute SKU of the training cluster
- C The train and evaluation split ratio
- D The max_total_trials sweep limit

Answer: D

The sweep limits setting caps the trials, concurrency, and time that the search may use. The compute SKU is the tempting setting to tune for a faster sweep, but it only changes the hardware of each trial. Evaluation splits and replica counts do not limit the search.

4. A _____ drift signal tracks how the distribution of a model's outputs changes compared with the validation period.

Answer: prediction (also accepted: prediction drift)

Prediction drift watches the output distribution, which can reveal a shift before ground truth labels arrive. Data drift is the tempting choice, but it watches inputs rather than outputs. Use both signals to see whether a change comes from the data or from the model's behavior.

5. Which two steps keep GenAI evaluation results reproducible? (Choose two.)

- A Re-sample a fresh test set for every evaluation run
- B Use the latest model version without recording which version was used
- C Version the test dataset and record the flow and model versions for each run
- D Clear the run history after each evaluation
- E Log the evaluator names and settings with each run

Answer: C, E

Versioning the dataset and recording the flow and model versions lets a run be repeated exactly, and logging evaluator settings keeps the scoring rules visible. Re-sampling the test set is the tempting way to avoid overfitting, but it breaks comparisons across runs. Clearing history removes the evidence that a comparison needs.

6. A team deploys a model to a managed online endpoint and must set how much capacity each deployment gets. Which settings control this?

- A instance_type and instance_count settings
- B The workspace description field used for documentation
- C The model's tags
- D The dataset version

Answer: A

The instance type and instance count in the deployment definition set the capacity and cost of each deployment. Model tags are the tempting thing to edit when labeling a release, but tags do not affect capacity. The workspace description and dataset version have no bearing on instance sizing.

7. Which artifact format lets an MLflow model deploy to Azure Machine Learning without custom inference code?
- A An MLflow tracking URI string
 - B MLflow autologging flags set
 - C An MLflow model package
 - D A raw pickle file on a datastore

Answer: C

An MLflow model package bundles the model, its signature, and its environment, so the platform can serve it without a custom scoring script. A raw pickle file is the tempting option because it holds the same weights, but it lacks the signature and environment the platform needs. A tracking URI only tells the client where to log runs.

8. Scoring a prompt flow only on the questions used while tuning the prompt gives an unbiased estimate of answer quality.
- A True
 - B False

Answer: B (False)

Evaluating on questions used during tuning overstates quality, so a held-out set is needed for an honest estimate. Reusing the tuning set is the tempting shortcut because it is already labeled, but it measures memory rather than generalization.

9. A team must keep the workspace storage and Key Vault reachable only from its private network. Which feature supports this?
- A Public blob access
 - B Shared access signatures with a long expiry
 - C Anonymous read on the storage container
 - D Private endpoints with network isolation

Answer: D

Private endpoints and network isolation keep traffic to the workspace resources inside the virtual network. Public blob access is the tempting setting when connectivity problems appear, but it widens exposure. Shared access signatures and anonymous reads both bypass network controls.

10. Which endpoint type suits a fraud check that must return a score within a few hundred milliseconds for each transaction?
- A Managed online endpoint
 - B Batch endpoint
 - C Scheduled pipeline job
 - D Notebook server

Answer: A

Managed online endpoints serve real-time requests with low latency. The batch endpoint is the tempting choice when many transactions must be scored, but it returns results asynchronously. Scheduled pipelines and notebooks are not serving endpoints.