

PRACTICE QUESTIONS

USAI

Certified Artificial Intelligence Engineer

SHARED BY [u/jollynebula1986](#) on aiexamcerts

EDITION	Free Set 2
QUESTIONS	10
SHARED	October 15, 2026
EXAM	USAI Certified Artificial Intelligence Engineer

Shared by [u/jollynebula1986](#) on aiexamcerts

10 practice questions for the USAII Certified Artificial Intelligence Engineer exam, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/usaii/certified-artificial-intelligence-engineer. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. Which pandas operation groups rows by a key and computes an aggregate for each group?

- A sort_values
- B drop_duplicates
- C groupby
- D rename

Answer: C

groupby splits rows into groups by a key so that an aggregate such as mean or count can be applied to each group. sort_values only reorders rows, and drop_duplicates removes repeated rows. The sort option is the tempting distractor, since sorting is often done just before grouping, but it does not compute group totals.

2. Python's garbage collector runs only when the program exits.

- A True
- B False

Answer: B (False)

CPython frees most objects through reference counting as soon as their counts reach zero, and a cyclic collector runs during execution to reclaim reference cycles. Memory is reclaimed while the program runs. The tempting belief is that cleanup is deferred until exit.

3. In hypothesis testing, what is the significance level alpha?

- A The probability that the alternative hypothesis is true
- B The sample size required for every test in the study
- C The power of the test to detect a real effect when one exists
- D The chance of rejecting a true null hypothesis

Answer: D

Alpha is the false positive rate you accept: the probability of rejecting the null hypothesis when it is true, chosen before the test is run. The power option is the tempting distractor, since power and alpha are both error-related, but power is the chance of detecting a real effect.

4. Bayes' theorem updates a prior probability using new _____ to produce a posterior probability.

Answer: evidence (also accepted: data)

Bayes' theorem combines the prior belief with the likelihood of the new evidence to produce an updated posterior, which then becomes the starting point for later updates. The word proof is a tempting alternative because it sounds rigorous, but the theorem works with evidence and likelihoods.

5. Which two statements about feature scaling are accurate? (Choose two.)

- A Scaling changes the labels of a classification target
- B Scaling is required before every decision tree is trained, or it fails
- C Standardization rescales features to zero mean and unit variance
- D Min-max scaling maps values into a fixed range such as zero to one
- E Scaling should be fitted on the test set to match the training set

Answer: C, D

Standardization centers features and scales them to unit variance, while min-max scaling maps them into a fixed range. Decision trees do not require scaling, and scalers should be fitted on training data only. The test-set option is the tempting distractor, since matching the test set sounds consistent, but it leaks information from the test data.

6. A logistic regression model outputs a probability for the positive class. Which function produces that output?

- A ReLU, which outputs zero for negative inputs
- B Tanh, which maps values into the range minus one to one
- C Sigmoid, which maps values into the range zero to one
- D A hard step function that outputs only zero or one

Answer: C

The sigmoid squashes any real-valued score into the interval from zero to one, which is why logistic regression uses it for a binary probability. ReLU and tanh are common in neural networks but do not map to a probability in this way, and a hard step gives no calibrated probability. Tanh is the tempting distractor, since it also squashes values, but its range is minus one to one.

7. A decision tree fits the training data perfectly but performs poorly on new data. What is the most likely issue?
- A Underfitting, because the tree is too shallow to capture the pattern
 - B The tree cannot handle categorical features at all in any library
 - C The learning rate is too small for the tree to converge properly during training
 - D Overfitting, because the tree memorized noise in the training data

Answer: D

A perfect training fit with poor test performance signals overfitting, where the tree has learned noise specific to the training set. Underfitting would show poor performance on both sets, and trees do handle categorical features. The underfitting option is the tempting distractor, since shallow models are a real cause of poor results, but it does not match perfect training accuracy.

8. K-means can produce different clusters on two runs with different random initializations.
- A True
 - B False

Answer: A (True)

K-means converges to a local optimum that depends on the starting centroids, which is why implementations run several initializations and keep the best result. The tempting belief is that k-means always finds the same global solution.

9. Training loss keeps falling while validation loss rises after a point. Which action is most appropriate?
- A Increase the number of epochs to see if the validation loss recovers
 - B Remove the validation set so the training loss looks better on the chart
 - C Stop training or add regularization, since the model overfits
 - D Double the learning rate to escape the validation curve quickly

Answer: C

Diverging training and validation loss is the classic overfitting signal, so early stopping or regularization is the right response. More epochs usually widen the gap, and removing the validation set hides the problem. The more-epochs option is the tempting distractor, since losses sometimes recover, but they rarely do so while the gap keeps widening.

10. Which hyperparameter most directly controls the size of each update step during gradient descent?
- A The number of output classes in the softmax layer
 - B The file format used to save the training checkpoint
 - C The number of cores on the training machine used
 - D The learning rate

Answer: D

The learning rate scales the gradient at each step, so it sets how far the weights move per update. The output class count defines the model's outputs, and file format or hardware cores do not change the optimization step. The class-count option is the tempting distractor, since it is a fixed design choice that shapes the output layer.