

PRACTICE QUESTIONS

USAI

Certified Artificial Intelligence Engineer

SHARED BY [u/breezy_dolphin_85](#) on aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
SHARED	October 15, 2026
EXAM	USAI Certified Artificial Intelligence Engineer

Shared by [u/breezy_dolphin_85](#) on aiexamcerts

10 practice questions for the USAI 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/usai/certified-artificial-intelligence-engineer. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. What does the pandas function `read_csv` return by default?
- A A NumPy array of the values, with no column labels attached
 - B A pandas DataFrame, with the first row used as column names
 - C A Python dictionary of raw strings, one entry for each line of the file
 - D A list of lists that contains every row, without any headers at all

Answer: B

`read_csv` parses the file into a DataFrame and by default treats the first row as the header. A NumPy array has no column labels, and a list of lists has no headers. The NumPy option is the tempting distractor, since DataFrames are often converted to arrays for modeling, but `read_csv` itself does not return one.

2. In pandas, missing numeric values are stored as _____, and functions such as `isna` and `dropna` detect and handle them.

Answer: NaN (also accepted: NA, np.nan)

Missing numeric values are stored as NaN, which stands for Not a Number, and `isna` and `dropna` are the standard ways to detect and remove them. Blank strings are not the same as missing values, so a column may need explicit cleaning. The word null is a tempting alternative, but it is a SQL term rather than the pandas representation.

3. Which probability distribution is most often used to model the number of events in a fixed time interval?
- A Normal distribution with a fixed mean
 - B Uniform distribution over the unit interval
 - C Poisson, for counts of events in a fixed interval
 - D Student t distribution with one degree of freedom

Answer: C

The Poisson distribution models counts of independent events in a fixed time or space interval when the average rate is known. The normal, uniform, and t distributions model other kinds of values. The normal distribution is the tempting distractor, because it is the most familiar, but it describes continuous values rather than event counts.

4. A correlation coefficient of zero means there is no relationship of any kind between two variables.
- A True
 - B False

Answer: B (False)

Pearson correlation measures only linear association, so a strong nonlinear relationship can produce a coefficient near zero. A zero value rules out a linear trend, not every relationship. The tempting belief is that zero correlation means independence.

5. Which two statements about k-nearest neighbors are accurate? (Choose two.)
- A It always produces a closed-form equation for the decision boundary
 - B It requires training a set of weights through gradient descent
 - C It predicts from the labels of the nearest training examples
 - D Feature scaling usually matters because distances drive the predictions
 - E It cannot be used for classification tasks

Answer: C, D

KNN predicts from the labels of the closest training points, and because distance drives the result, unscaled features can dominate it. KNN has no weight-training step, and classification is a standard use. The weight-training option is the tempting distractor, since most supervised models learn weights, but KNN stores the examples instead.

6. What is the role of a validation set?
- A To provide the final unbiased estimate after all decisions are made
 - B To tune hyperparameters and compare candidate models
 - C To store the model weights for deployment to production servers
 - D To replace the need for feature engineering in the pipeline

Answer: B

The validation set guides choices such as hyperparameters and model selection, which keeps the test set untouched for an honest final estimate. The final-estimate option describes the test set, which is the tempting distractor because the two sets are easy to confuse. Storing weights and skipping feature work are unrelated to the validation role.

7. What does k-means clustering require the analyst to choose in advance?

- A The exact label for every record in the dataset before fitting
- B The number of clusters, k
- C A target column that the model will predict during training
- D The number of hidden layers in a neural network architecture

Answer: B

K-means needs the number of clusters specified up front, and it then assigns points to the nearest centroid. Labels and target columns are not inputs because the method is unsupervised, and hidden layers belong to neural networks. The labels option is the tempting distractor, since clusters are often checked against labels, but labels are not an input to the algorithm.

8. Creating a new input such as a ratio of two columns, computed from raw data with domain knowledge, is called feature _____.

Answer: engineering

Feature engineering creates new inputs from raw data, and it often improves a model more than changing the algorithm. The derived ratio must be computed the same way at prediction time. Selection is a tempting alternative, but it chooses among existing features rather than creating new ones.

9. What does batch normalization do during training?

- A It normalizes each mini-batch's activations to stabilize training
- B It removes all bias terms from the network permanently after training
- C It converts image inputs into batches of text tokens for processing
- D It saves a full copy of the model after every epoch to disk storage

Answer: A

Batch normalization standardizes activations within each mini-batch and applies learned scale and shift, which stabilizes training and often allows higher learning rates. It does not remove biases, convert data types, or checkpoint models. The checkpoint option is the tempting distractor, since saving copies is a real training practice, but it is not what batch normalization does.

10. Why are residual connections used in deep networks?

- A They remove the need for any activation function in the network
- B They let gradients flow through very deep stacks
- C They turn a classifier into a clustering model without relabeling
- D They guarantee that the network will never overfit the training data

Answer: B

A skip connection adds the input of a block to its output, which gives gradients a direct path and eases vanishing gradients in deep networks. It does not remove activations, change the task, or prevent overfitting. The overfitting claim is the tempting distractor, since deeper networks can seem more reliable, but skip connections do not guarantee generalization.