

PRACTICE QUESTIONS

CertNexus

AIP-210

Certified Artificial Intelligence Practitioner (CAIP)

SHARED BY [u/snappyrouter1993](#) on aiexamcerts

EDITION Free Set 2

QUESTIONS 10

SHARED October 15, 2026

EXAM CertNexus AIP-210, Certified Artificial Intelligence Practitioner (CAIP)

Shared by [u/snappyrouter1993](#) on aiexamcerts

10 practice questions for the CertNexus AIP-210 exam, Certified Artificial Intelligence Practitioner (CAIP), each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/certnexus/aip-210. Spotted a mistake? Email support@aiexamcerts.com with the exam code AIP-210 and the question number.

1. Which practice best prevents data leakage when evaluating a model with a scaler and an imputer?
 - A Fit the scaler and imputer on the training split only, then apply them to the test split
 - B Fit the preprocessing steps on the full dataset before splitting so the statistics match
 - C Remove the test split once the model has been trained once
 - D Shuffle the rows after the model has been evaluated

Answer: A

Fitting preprocessing on the training split and only applying it to the test split keeps test-set statistics out of training. Fitting on the full dataset before splitting is the most tempting mistake, because it leaks information from the test rows into the model. Removing the test split or shuffling after evaluation does not address leakage at all.

2. Early stopping halts training when validation performance has not improved for a set number of epochs.
 - A True
 - B False

Answer: A (True)

Early stopping monitors a validation metric and stops training, or restores the best checkpoint, once the metric has failed to improve for a chosen patience period. It is a regularization method that prevents the model from drifting into overfitting. The statement is True.

3. Which is the best first step before building an AI solution for a customer support team?
 - A Choose the largest available model and deploy it to every customer channel right away
 - B Define the problem, the success metric, and the data, then test a simple baseline
 - C Buy GPU hardware before any use case has been defined
 - D Promise full automation of every support ticket at launch

Answer: B

Clarifying the problem, success criteria, and data prevents building the wrong thing and shows whether AI is needed at all. Choosing the largest model is tempting, but it skips scoping and cost control. Buying hardware first and promising full automation are common causes of failed projects.

4. The metric that takes the square root of the average squared difference between predictions and actual values is called root _____ error.

Answer: mean squared

Root mean squared error is expressed in the same units as the target, which makes it easier to interpret than MSE. The squaring penalizes large errors more heavily. Mean absolute error is a related measure that does not square the residuals.

5. Which two statements about the relationship between artificial intelligence, machine learning, and deep learning are correct? (Choose two.)
- A Deep learning is a subset of machine learning that uses multi-layer neural networks
 - B Machine learning is a subset of artificial intelligence that learns patterns from data
 - C Artificial intelligence is a subset of deep learning that runs on GPUs
 - D Machine learning always requires hand-written if-then rules
 - E Deep learning cannot be used for image recognition

Answer: A, B

Each field sits inside the one before it: AI is the broadest, machine learning learns from data within AI, and deep learning uses multi-layer neural networks within machine learning. The option claiming machine learning requires hand-written rules is the most tempting wrong answer, because rule-based systems are an early AI approach, but machine learning is defined by learning rules from data. Deep learning is widely used for image recognition, so the last option is plainly false.

6. A spam filter flags 40 emails, and 30 of them really are spam. What is the precision of the filter for the spam class?
- A 0.30
 - B 0.40
 - C 0.75
 - D 0.70

Answer: C

Precision is true positives divided by all predicted positives, here 30 out of 40, which equals 0.75. Confusing precision with recall is the most tempting error, but recall needs the total number of real spam emails, which the scenario does not give. The value 0.30 is not the right ratio for this data.

7. Why is one-hot encoding used for a categorical feature such as region, with values North, South, East, and West?
- A It converts the categories into one integer that ranks regions by sales
 - B It removes the feature because models cannot use categorical data
 - C It groups the regions into two numeric buckets based on alphabetical order
 - D It creates one 0/1 column per category, so no order is assumed

Answer: D

One-hot encoding creates a 0 or 1 column for each category, which avoids the false assumption that North is less than South. Ordinal integer codes are the most tempting alternative, but they imply a ranking that does not exist here. Removing the feature throws away useful information.

8. Retraining a model on newer data always improves its performance in production.
- A True
 - B False

Answer: B (False)

Retraining helps when the data has shifted, but new data can be mislabeled, unrepresentative, or introduce leakage, so every retrained model must be evaluated before it is promoted. The statement overgeneralizes and is False.

9. A team trains a convolutional network on a small labeled image set and sees strong overfitting. Which technique most directly reduces this?
- A Data augmentation such as random crops and horizontal flips
 - B Doubling the number of training epochs with no other change
 - C Removing dropout and batch normalization layers
 - D Raising the learning rate while training on the same images

Answer: A

Augmentation creates varied versions of the existing images, which encourages the network to learn invariances instead of memorizing pixels. Training for more epochs usually worsens overfitting on small data, which is the most tempting misstep. Removing regularization layers and raising the learning rate add no new variation to the data.

10. Missing a positive case costs far more than a false alarm, as in screening for a serious disease. Which metric should the team prioritize?
- A Specificity
 - B Recall
 - C Mean absolute error
 - D Silhouette score

Answer: B

Recall measures the share of actual positives that the model finds, so it directly penalizes missed cases.

Specificity focuses on correctly identifying negatives, which is the tempting but opposite priority. Mean absolute error is a regression metric and the silhouette score evaluates clustering.