

PRACTICE QUESTIONS

ARTiBA

Artificial Intelligence Engineer

SHARED BY [u/stellarrobin2004](#) on aiexamcerts

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

Shared by [u/stellarrobin2004](#) on aiexamcerts

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

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

1. Which term describes a model that performs well on its training data but poorly on unseen data?
- A Underfitting
 - B Overfitting
 - C Calibration drift
 - D Regularization

Answer: B

Overfitting means the model has learned noise and quirks specific to the training set, so it generalizes poorly. Underfitting is the tempting confusion, but an underfit model performs poorly on both training and unseen data. Regularization is a remedy for overfitting rather than the problem itself.

2. In a train, validation, and test split, the _____ set is used only once, at the end, to estimate performance on unseen data.

Answer: test (also accepted: test set, held-out test set)

The test set is held back until the model and its hyperparameters are finalized, so its score is an unbiased estimate of generalization. The validation set is the one reused during tuning, which is why it cannot serve as the final estimate.

3. A product team needs clickstream events to be available for analysis within seconds of being generated. Which architecture best fits this requirement?
- A A streaming message broker with consumers that write to an analytical store
 - B A nightly batch export to a flat file
 - C Manual CSV uploads by analysts each Friday, reviewed by hand before anyone sees the data
 - D A weekly full database dump

Answer: A

A streaming broker buffers events as they arrive and lets consumers write them onward continuously, meeting the seconds-level requirement. The nightly export is the tempting choice because it is simpler to build, but its latency is measured in hours. Manual uploads and weekly dumps fail the timeliness requirement outright.

4. Imputing missing numeric values with the mean is generally more robust to extreme outliers than imputing with the median.
- A True
 - B False

Answer: B (False)

The statement is false. A few extreme values pull the mean toward them, while the median stays near the middle of the distribution, so median imputation is usually the more robust choice for skewed numeric columns.

5. Which two statements about ROC curves for a binary classifier are correct? (Choose two.)
- A It plots true positive rate against false positive rate across decision thresholds
 - B A random classifier traces the diagonal from (0,0) to (1,1), with an area of 0.5
 - C It plots precision against recall
 - D A perfect classifier has an area under the curve of 0
 - E Raising the threshold always increases recall

Answer: A, B

The ROC curve traces true positive rate against false positive rate as the threshold moves, and a random guesser lies on the diagonal with an area of 0.5. Plotting precision against recall describes the precision-recall curve, which is the tempting confusion. A perfect classifier has an area of 1, and raising the threshold usually lowers recall.

6. A team is building a regression model that predicts house prices in dollars, and large errors are much more costly to the business than small ones. Which metric is most appropriate?
- A Accuracy
 - B F1 score
 - C Mean absolute error
 - D Root mean squared error

Answer: D

Root mean squared error is expressed in the target's units and squares each error, so large misses weigh more heavily in the score. Mean absolute error is the tempting alternative, but it treats all errors linearly and penalizes large misses less. Accuracy and F1 score are classification metrics and do not fit a continuous price target.

7. Which loss function is most commonly paired with a softmax output layer for multi-class classification?

- A Mean squared error
- B Mean absolute error
- C Binary cross-entropy
- D Categorical cross-entropy

Answer: D

Categorical cross-entropy compares the predicted class probability distribution from softmax with the one-hot target. Binary cross-entropy is the tempting choice, but it is intended for a single output that separates two classes. Mean squared error and mean absolute error are regression losses.

8. A recurrent neural network processes a sequence by passing a _____ state from one time step to the next.

Answer: hidden (also accepted: hidden state)

The hidden state summarizes what the network has seen so far and is updated at each step, which is what gives recurrent models their memory of order. This step-by-step dependence is also why long sequences can be slow to train.

9. Which term describes a numeric vector representation of text in which semantically similar items sit near one another?

- A Embedding
- B Tokenization
- C Checkpoint
- D Gradient

Answer: A

Embeddings encode meaning as dense vectors, so similarity search can find passages with related content even when the wording differs. Tokenization is the tempting confusion because tokens are how text enters a model, but tokens are discrete pieces of text rather than vectors. A checkpoint is a saved model state and a gradient is a training signal.

10. Which metric is commonly used to compare the similarity of two text embedding vectors?

- A Learning rate
- B Cosine similarity
- C Dropout rate
- D Batch size

Answer: B

Cosine similarity compares the direction of two vectors and ignores their length, which suits embeddings whose magnitudes carry little meaning. The learning rate, dropout rate, and batch size are training settings rather than similarity measures, so they are easy to rule out once the question is read carefully.