

PRACTICE QUESTIONS

PMI Certified Professional in Managing AI

SHARED BY [u/Jolly-Cluster-5776](#) on aiexamcerts

EDITION	Free Set 2
QUESTIONS	10
SHARED	October 15, 2026
EXAM	PMI Certified Professional in Managing AI

Shared by [u/Jolly-Cluster-5776](#) on aiexamcerts

10 practice questions for the PMI Certified Professional in Managing AI exam, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/pmi/pmi-certified-professional-in-managing-ai. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. Which success criterion is most useful in the business understanding phase?

- A Training loss at the final epoch
- B The number of parameters in the model
- C Reduction in manual review hours at an agreed accuracy threshold
- D The count of GPUs used during training

Answer: C

A success criterion should connect model performance to a business outcome, such as review hours saved while holding an agreed accuracy level. Training loss is a technical signal that means little to the business. Parameter count is the tempting distractor because it sounds like a measure of capability, but it says nothing about results.

2. A model trained on data from only one customer segment is safe to deploy to every segment.

- A True
- B False

Answer: B (False)

A model learns the patterns of its training population, so applying it to segments that were absent from training is a coverage gap. Data understanding should check whether the training population matches the deployment population. The statement is false.

3. A medical triage model has high overall accuracy but misses many urgent cases. Which metric should be examined to reveal the miss rate on urgent cases?

- A Precision
- B Specificity
- C Recall
- D Mean absolute error

Answer: C

Recall measures the share of true urgent cases that the model catches, so low recall exposes the misses that overall accuracy hides. Precision is the tempting distractor because it is also a common clinical metric, but it measures how many flagged cases are correct rather than how many urgent cases were missed. Mean absolute error is a regression metric and does not describe misses on a class.

4. The CPMAI methodology is _____, which means the team returns to earlier phases whenever new data or results change the picture.

Answer: iterative

CPMAI is built around iteration, so teams return to business or data understanding when a model result reveals a gap. A strictly linear plan is the wrong mental model for AI work. Following a waterfall sequence is the tempting alternative, but each phase depends on what was learned in the one before it.

5. Which two items belong in a production handoff for a deployed model? (Choose two.)

- A The developer's personal notebook with exploratory plots
- B A plan to delete the previous model version to save storage
- C A runbook with monitoring thresholds and a named owner who responds to alerts
- D A promise that the model will never change after handoff
- E Versioned model artifacts with the training data lineage recorded

Answer: C, E

Operations needs a runbook with alert thresholds and a named owner, plus versioned artifacts that trace back to their training data, so issues can be diagnosed and reproduced. A personal notebook is the tempting distractor, since it holds the real work, but it is neither reproducible nor supported. Deleting the previous version removes the rollback path the team will need during an incident.

6. Which document most directly records a model's intended use, limitations, and evaluation results for stakeholders?

- A A cost report from the cloud provider
- B A model card
- C The source code commit log of the project
- D The network access policy for the training cluster

Answer: B

A model card summarizes intended use, training data, evaluation results across groups, and known limitations, which is exactly what reviewers and stakeholders need. The commit log is the tempting distractor because it records every change to the code, but it does not explain how the model should be used or where it fails.

7. Which approach keeps a categorical feature with thousands of unique values from creating an impractically wide feature set?
- A Grouping rare categories and fitting any target-aware encoding on training folds only
 - B One-hot encoding every value with no grouping, which creates a column for each unique category
 - C Converting each category to a random integer and ignoring its meaning
 - D Removing all categorical features from every model

Answer: A

Grouping rare levels and fitting any target-aware encoding only on training folds keeps the feature space manageable without leaking target information. One-hot encoding every value is the tempting distractor, since it is the default in many libraries, but with high cardinality it produces a very wide and sparse matrix. Random integer codes add false ordering to the feature.

8. A model with a high test score on a random sample is guaranteed to perform well on rare but critical cases.
- A True
 - B False

Answer: B (False)

Aggregate test scores are dominated by common cases, so rare or critical cases need their own slices and metrics. A sample that under-represents them makes the overall score misleading. The statement is false.

9. A team wants to know whether a customer-support dataset is complete enough for a model. Which check is most direct?
- A Counting the columns in the spreadsheet and confirming that each column has a header name
 - B Profiling field completeness and the share of records with missing critical values
 - C Checking the file's creation date
 - D Comparing the file size with the size of last month's extract

Answer: B

Completeness profiling shows which critical fields are missing and how often, which determines whether the data can support the model. Counting columns is the tempting distractor, since it looks like a data check, but it says nothing about whether the values are actually present. Neither the creation date nor a file size comparison measures completeness.

10. Which practice best addresses the risk that a model's errors fall unevenly on one demographic group?
- A Evaluating only overall accuracy
 - B Removing the protected attribute column and never checking outcomes by group afterward
 - C Averaging results across all groups so that no single group stands out
 - D Measuring performance by group and setting acceptance thresholds for each group

Answer: D

Disaggregated evaluation shows whether errors concentrate on a group, and acceptance thresholds turn that finding into a release decision. Removing the protected attribute is the tempting distractor, since it sounds like a fix, but proxy features can carry the same bias and outcomes still have to be checked.