

PRACTICE QUESTIONS

UiPath Specialized AI Professional v1.0

SHARED BY [u/glad_comet_17](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 15, 2026

EXAM UiPath Specialized AI Professional v1.0

Shared by u/glad_comet_17 on aiexamcerts

10 practice questions for the UiPath Specialized AI Professional v1.0 exam, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/uipath/uipath-specialized-ai-professional-v1-0. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. Which statement about embeddings is correct?

- A Embeddings are encrypted copies of documents stored in the index, with the original text kept in a separate, readable column
- B Embeddings are the final answers returned by a language model
- C Embeddings count the exact number of words in each passage
- D Embeddings map text to numeric vectors so that passages with similar meaning sit close together

Answer: D

An embedding places text in a vector space where distance reflects similarity of meaning, which powers semantic search. Encrypted copies are the tempting distractor, since indexes do store protected content, but an embedding is a vector, not a copy. Final answers and exact word counts are not what embeddings represent.

2. The numeric representation of text that is used to find passages with similar meaning is called an _____.

Answer: embedding

An embedding is a vector that captures meaning, and similar texts produce nearby embeddings. An index is the tempting alternative, since the vectors are stored in one, but the index is the store, not the representation. The representation itself is the embedding.

3. What does a human escalation path in an agent design usually handle?

- A Every routine case, so humans always decide and the agent never finishes a task alone
- B Only the formatting of the final report, with nothing else passed to a person
- C Cases where the agent is uncertain, the action is high-risk, or policy requires an approval
- D Only the robot restart schedule and the license renewals each year

Answer: C

Escalation routes uncertain, risky or policy-bound cases to a person, while routine work runs automatically. Reviewing every case is the tempting distractor, because it sounds safe, but it removes most of the benefit. Formatting, restart schedules and renewals are not escalation triggers.

4. A test set should be used repeatedly to tune hyperparameters so that the final score improves.
- A True
 - B False

Answer: B (False)

Repeated tuning on the test set leaks it into the model choice and inflates the reported score. The tempting error is assuming the test set is just another validation set. A separate validation set is used for tuning.

5. Which two practices improve retrieval quality in a grounded pipeline? (Choose two.)
- A Indexing every file regardless of whether it is current
 - B Adding metadata filters for document type and date
 - C Returning the top fifty passages to the model for every question
 - D Removing headings and titles so that passages look uniform
 - E Tuning chunk size so that passages hold complete ideas

Answer: B, E

Metadata filters and sensible chunk sizes both improve the evidence the model receives. Indexing outdated files is the tempting distractor, because more content seems helpful, but it adds stale evidence. Returning very many passages and stripping headings both reduce the precision of retrieval.

6. A classifier assigns a document to a type. Which later step extracts its fields?
- A The Orchestrator retry policy for the queue, which runs whenever a document is classified
 - B The license allocation report for the tenant
 - C The Studio package signature for the project
 - D An extractor configured for that document type

Answer: D

Once the type is known, the matching extractor reads the fields defined for that type in the taxonomy. The retry policy is the tempting distractor, because it is also a processing setting, but it governs failures rather than field reading. Licenses and package signatures are unrelated to extraction.

7. A dataset has about 2% positive cases. Which metric is most misleading to report on its own?
- A Precision, because it is the single best number to report for rare classes
 - B Accuracy, because a model that predicts negative every time scores about 98%
 - C Recall, because it counts the positive cases the model finds among all the actual positives
 - D The F1 score, because it combines precision and recall in one value

Answer: B

With 2% positives, always predicting negative gets about 98% accuracy while finding no positives at all. Precision is the tempting distractor, because it is also sometimes unstable on rare classes, but it does not reward ignoring the positives. Recall and F1 both expose misses on the rare class.

8. The harmonic mean of precision and recall is called the _____ score.

Answer: F1 (also accepted: F1-score, F-1)

The F1 score balances precision and recall, so it drops when either is weak. Accuracy is the tempting alternative, because it is the most familiar measure, but accuracy can look high on imbalanced fields. F1 is the usual single number when both errors matter.

9. What is the purpose of a confidence threshold on an extraction result?

- A It sets the maximum number of pages a document can have
- B It controls how many robots can run in parallel on a queue
- C It decides whether a result is accepted automatically or sent for human validation
- D It determines the OCR engine used to read each page

Answer: C

A threshold separates results the pipeline can trust from those that need a person to check them. The page limit is the tempting distractor, since document size rules are also configured, but the threshold does not set page counts. Robot parallelism and the OCR engine choice are unrelated settings.

10. In retrieval grounding, why are long documents split into chunks before indexing?

- A The index cannot store any file longer than one page, so every page is its own file
- B Chunks let the model skip the retrieval step entirely in the pipeline
- C Chunking converts the documents into images so that OCR can run on every page
- D Smaller passages can be matched to a question more precisely, and they fit within the model's context window

Answer: D

Passages that cover one idea match questions more precisely and leave room in the prompt for several of them. The one-page limit is the tempting distractor, because page limits do exist in some tools, but they are not the reason for chunking. Skipping retrieval and converting to images are not purposes of chunking.