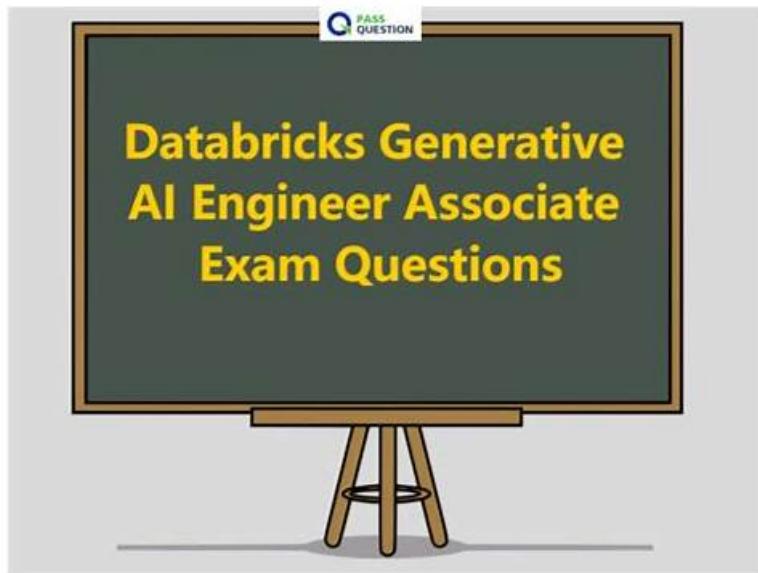


# Databricks Databricks-Generative-AI-Engineer-Associate Exam Questions: Attain Your Professional Career Targets [2026]



BTW, DOWNLOAD part of Exam4Labs Databricks-Generative-AI-Engineer-Associate dumps from Cloud Storage:  
[https://drive.google.com/open?id=1QAgv6nDBk5furnj\\_W1NDlCY\\_MiEQwjgAP](https://drive.google.com/open?id=1QAgv6nDBk5furnj_W1NDlCY_MiEQwjgAP)

Obtaining the Databricks-Generative-AI-Engineer-Associate certificate will make your colleagues and supervisors stand out for you, because it represents Databricks-Generative-AI-Engineer-Associate your professional skills. At the same time, it will also give you more opportunities for promotion and job-hopping. The Databricks-Generative-AI-Engineer-Associate latest exam dumps have different classifications for different qualification examinations, which can enable students to choose their own learning mode for themselves according to the actual needs of users. The Databricks-Generative-AI-Engineer-Associate Exam Question offer a variety of learning modes for users to choose from, which can be used for multiple clients of computers and mobile phones to study online, as well as to print and print data for offline consolidation.

## Databricks Databricks-Generative-AI-Engineer-Associate Exam Syllabus Topics:

Topic	Details
Topic 1	<ul style="list-style-type: none"><li>Evaluation and Monitoring: This topic is all about selecting an LLM choice and key metrics. Moreover, Generative AI Engineers learn about evaluating model performance. Lastly, the topic includes sub-topics about inference logging and usage of Databricks features.</li></ul>
Topic 2	<ul style="list-style-type: none"><li>Data Preparation: Generative AI Engineers covers a chunking strategy for a given document structure and model constraints. The topic also focuses on filter extraneous content in source documents. Lastly, Generative AI Engineers also learn about extracting document content from provided source data and format.</li></ul>
Topic 3	<ul style="list-style-type: none"><li>Design Applications: The topic focuses on designing a prompt that elicits a specifically formatted response. It also focuses on selecting model tasks to accomplish a given business requirement. Lastly, the topic covers chain components for a desired model input and output.</li></ul>
Topic 4	<ul style="list-style-type: none"><li>Application Development: In this topic, Generative AI Engineers learn about tools needed to extract data, Langchain</li><li>similar tools, and assessing responses to identify common issues. Moreover, the topic includes questions about adjusting an LLM's response, LLM guardrails, and the best LLM based on the attributes of the application.</li></ul>

Topic 5	<ul style="list-style-type: none"><li>• Governance: Generative AI Engineers who take the exam get knowledge about masking techniques, guardrail techniques, and legal</li><li>• licensing requirements in this topic.</li></ul>
---------	---

>> Reliable Databricks-Generative-AI-Engineer-Associate Exam Price <<

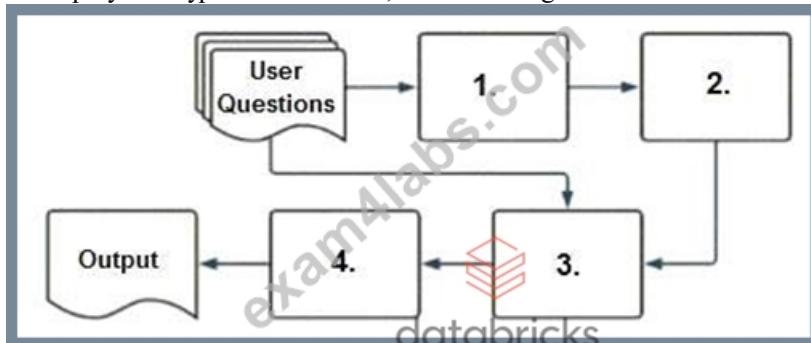
## **Major Formats of Databricks Databricks-Generative-AI-Engineer-Associate Exam Questions**

There are plenty of platforms that have been offering Databricks Certified Generative AI Engineer Associate Databricks-Generative-AI-Engineer-Associate exam practice questions. You have to be vigilant and choose the reliable and trusted platform for Databricks Certified Generative AI Engineer Associate Databricks-Generative-AI-Engineer-Associate exam preparation and the best platform is Exam4Labs. On this platform, you will get the valid, updated, and Databricks Certified Generative AI Engineer Associate exam expert-verified exam questions. Databricks Certified Generative AI Engineer Associate Questions are real and error-free questions that will surely repeat in the upcoming Databricks Certified Generative AI Engineer Associate exam and you can easily pass the final Databricks Certified Generative AI Engineer Associate Databricks-Generative-AI-Engineer-Associate Exam even with good scores.

## **Databricks Certified Generative AI Engineer Associate Sample Questions (Q52-Q57):**

## NEW QUESTION # 52

A company has a typical RAG-enabled, customer-facing chatbot on its website.



Select the correct sequence of components a user's question will go through before the final output is returned. Use the diagram above for reference.

- A. 1.embedding model, 2.vector search, 3.context-augmented prompt, 4.response-generating LLM
- B. 1.response-generating LLM, 2.vector search, 3.context-augmented prompt, 4.embedding model
- C. 1.response-generating LLM, 2.context-augmented prompt, 3.vector search, 4.embedding model
- D. 1.context-augmented prompt, 2.vector search, 3.embedding model, 4.response-generating LLM

Answer: A

### Explanation:

To understand how a typical RAG-enabled customer-facing chatbot processes a user's question, let's go through the correct sequence as depicted in the diagram and explained in option A:

\* Embedding Model (1): The first step involves the user's question being processed through an embedding model. This model converts the text into a vector format that numerically represents the text. This step is essential for allowing the subsequent vector search to operate effectively.

\* Vector Search (2): The vectors generated by the embedding model are then used in a vector search mechanism. This search identifies the most relevant documents or previously answered questions that are stored in a vector format in a database.

\* Context-Augmented Prompt (3): The information retrieved from the vector search is used to create a context-augmented prompt. This step involves enhancing the basic user query with additional relevant information gathered to ensure the generated response is as accurate and informative as possible.

\* Response-Generating LLM (4): Finally, the context-augmented prompt is fed into a response-generating large language model (LLM). This LLM uses the prompt to generate a coherent and contextually appropriate answer, which is then delivered as the final

output to the user.

Why Other Options Are Less Suitable:

\* B, C, D: These options suggest incorrect sequences that do not align with how a RAG system typically processes queries. They misplaced the role of embedding models, vector search, and response generation in an order that would not facilitate effective information retrieval and response generation.

Thus, the correct sequence is embedding model, vector search, context-augmented prompt, response-generating LLM, which is option A.

### NEW QUESTION # 53

A Generative AI Engineer has developed an LLM application to answer questions about internal company policies. The Generative AI Engineer must ensure that the application doesn't hallucinate or leak confidential data.

Which approach should NOT be used to mitigate hallucination or confidential data leakage?

- A. Use a strong system prompt to ensure the model aligns with your needs.
- B. Limit the data available based on the user's access level
- C. Add guardrails to filter outputs from the LLM before it is shown to the user
- D. **Fine-tune the model on your data, hoping it will learn what is appropriate and not**

**Answer: D**

Explanation:

When addressing concerns of hallucination and data leakage in an LLM application for internal company policies, fine-tuning the model on internal data with the hope it learns data boundaries can be problematic:

\* Risk of Data Leakage: Fine-tuning on sensitive or confidential data does not guarantee that the model will not inadvertently include or reference this data in its outputs. There's a risk of overfitting to the specific data details, which might lead to unintended leakage.

\* Hallucination: Fine-tuning does not necessarily mitigate the model's tendency to hallucinate; in fact, it might exacerbate it if the training data is not comprehensive or representative of all potential queries.

Better Approaches:

\* A, C, and D involve setting up operational safeguards and constraints that directly address data leakage and ensure responses are aligned with specific user needs and security levels.

Fine-tuning lacks the targeted control needed for such sensitive applications and can introduce new risks, making it an unsuitable approach in this context.

### NEW QUESTION # 54

A Generative AI Engineer is deciding between using LSH (Locality Sensitive Hashing) and HNSW (Hierarchical Navigable Small World) for indexing their vector database. Their top priority is semantic accuracy. Which approach should the Generative AI Engineer use to evaluate these two techniques?

- A. Compare the Levenshtein distances of returned results against a representative sample of test inputs
- B. Compare the Bilingual Evaluation Understudy (BLEU) scores of returned results for a representative sample of test inputs
- C. **Compare the cosine similarities of the embeddings of returned results against those of a representative sample of test inputs**
- D. Compare the Recall-Oriented-Understudy for Gisting Evaluation (ROUGE) scores of returned results for a representative sample of test inputs

**Answer: C**

Explanation:

The task is to choose between LSH and HNSW for a vector database index, prioritizing semantic accuracy.

The evaluation must assess how well each method retrieves semantically relevant results. Let's evaluate the options.

\* Option A: Compare the cosine similarities of the embeddings of returned results against those of a representative sample of test inputs

\* Cosine similarity measures semantic closeness between vectors, directly assessing retrieval accuracy in a vector database. Comparing returned results' embeddings to test inputs' embeddings evaluates how well LSH or HNSW preserves semantic relationships, aligning with the priority.

\* Databricks Reference: "Cosine similarity is a standard metric for evaluating vector search accuracy" ("Databricks Vector Search Documentation," 2023).

\* Option B: Compare the Bilingual Evaluation Understudy (BLEU) scores of returned results for a representative sample of test inputs

\* BLEU evaluates text generation (e.g., translations), not vector retrieval accuracy. It's irrelevant for indexing performance.

- \* Databricks Reference: "BLEU applies to generative tasks, not retrieval" ("Generative AI Cookbook").
- \* Option C: Compare the Recall-Oriented-Understudy for Gisting Evaluation (ROUGE) scores of returned results for a representative sample of test inputs
- \* ROUGE is for summarization evaluation, not vector search. It doesn't measure semantic accuracy in retrieval.
- \* Databricks Reference: "ROUGE is unsuited for vector database evaluation" ("Building LLM Applications with Databricks").
- \* Option D: Compare the Levenshtein distances of returned results against a representative sample of test inputs
- \* Levenshtein distance measures string edit distance, not semantic similarity in embeddings. It's inappropriate for vector-based retrieval.

\* Databricks Reference: No specific support for Levenshtein in vector search contexts.

Conclusion: Option A (cosine similarity) is the correct approach, directly evaluating semantic accuracy in vector retrieval, as recommended by Databricks for Vector Search assessments.

## NEW QUESTION # 55

A Generative AI Engineer is creating an agent-based LLM system for their favorite monster truck team. The system can answer text based questions about the monster truck team, lookup event dates via an API call, or query tables on the team's latest standings. How could the Generative AI Engineer best design these capabilities into their system?

- A. Ingest PDF documents about the monster truck team into a vector store and query it in a RAG architecture.
- B. Build a system prompt with all possible event dates and table information in the system prompt. Use a RAG architecture to lookup generic text questions and otherwise leverage the information in the system prompt.
- **C. Write a system prompt for the agent listing available tools and bundle it into an agent system that runs a number of calls to solve a query.**
- D. Instruct the LLM to respond with "RAG", "API", or "TABLE" depending on the query, then use text parsing and conditional statements to resolve the query.

### Answer: C

Explanation:

In this scenario, the Generative AI Engineer needs to design a system that can handle different types of queries about the monster truck team. The queries may involve text-based information, API lookups for event dates, or table queries for standings. The best solution is to implement an agent-based system.

Here's how option B works, and why it's the most appropriate answer:

- \* System Design Using Agent-Based Model: In modern agent-based LLM systems, you can design a system where the LLM (Large Language Model) acts as a central orchestrator. The model can "decide" which tools to use based on the query. These tools can include API calls, table lookups, or natural language searches. The system should contain a system prompt that informs the LLM about the available tools.
- \* System Prompt Listing Tools: By creating a well-crafted system prompt, the LLM knows which tools are at its disposal. For instance, one tool may query an external API for event dates, another might look up standings in a database, and a third may involve searching a vector database for general text-based information. The agent will be responsible for calling the appropriate tool depending on the query.
- \* Agent Orchestration of Calls: The agent system is designed to execute a series of steps based on the incoming query. If a user asks for the next event date, the system will recognize this as a task that requires an API call. If the user asks about standings, the agent might query the appropriate table in the database. For text-based questions, it may call a search function over ingested data. The agent orchestrates this entire process, ensuring the LLM makes calls to the right resources dynamically.

\* Generative AI Tools and Context: This is a standard architecture for integrating multiple functionalities into a system where each query requires different actions. The core design in option B is efficient because it keeps the system modular and dynamic by leveraging tools rather than overloading the LLM with static information in a system prompt (like option D).

\* Why Other Options Are Less Suitable:

\* A (RAG Architecture): While relevant, simply ingesting PDFs into a vector store only helps with text-based retrieval. It wouldn't help with API lookups or table queries.

\* C (Conditional Logic with RAG/API/TABLE): Although this approach works, it relies heavily on manual text parsing and might introduce complexity when scaling the system.

\* D (System Prompt with Event Dates and Standings): Hardcoding dates and table information into a system prompt isn't scalable. As the standings or events change, the system would need constant updating, making it inefficient.

By bundling multiple tools into a single agent-based system (as in option B), the Generative AI Engineer can best handle the diverse requirements of this system.

## NEW QUESTION # 56

A Generative AI Engineer has built an LLM-based system that will automatically translate user text between two languages. They now want to benchmark multiple LLM's on this task and pick the best one. They have an evaluation set with known high quality translation examples. They want to evaluate each LLM using the evaluation set with a performant metric. Which metric should they choose for this evaluation?

- A. BLEU metric
- B. ROUGE metric
- C. NDCG metric
- D. RECALL metric

#### Answer: A

Explanation:

The task is to benchmark LLMs for text translation using an evaluation set with known high-quality examples, requiring a performant metric. Let's evaluate the options.

\* Option A: ROUGE metric

\* ROUGE (Recall-Oriented Understudy for Gisting Evaluation) measures overlap between generated and reference texts, primarily for summarization. It's less suited for translation, where precision and word order matter more.

\* Databricks Reference: "ROUGE is commonly used for summarization, not translation evaluation" ("Generative AI Cookbook," 2023).

\* Option B: BLEU metric

\* BLEU (Bilingual Evaluation Understudy) evaluates translation quality by comparing n-gram overlap with reference translations, accounting for precision and brevity. It's widely used, performant, and appropriate for this task.

\* Databricks Reference: "BLEU is a standard metric for evaluating machine translation, balancing accuracy and efficiency" ("Building LLM Applications with Databricks").

\* Option C: NDCG metric

\* NDCG (Normalized Discounted Cumulative Gain) assesses ranking quality, not text generation.

It's irrelevant for translation evaluation.

\* Databricks Reference: "NDCG is suited for ranking tasks, not generative output scoring" ("Databricks Generative AI Engineer Guide").

\* Option D: RECALL metric

\* Recall measures retrieved relevant items but doesn't evaluate translation quality (e.g., fluency, correctness). It's incomplete for this use case.

\* Databricks Reference: No specific extract, but recall alone lacks the granularity of BLEU for text generation tasks.

Conclusion: Option B (BLEU) is the best metric for translation evaluation, offering a performant and standard approach, as endorsed by Databricks' guidance on generative tasks.

#### NEW QUESTION # 57

.....

Our evaluation system for Databricks-Generative-AI-Engineer-Associate test material is smart and very powerful. First of all, our researchers have made great efforts to ensure that the data scoring system of our Databricks-Generative-AI-Engineer-Associate test questions can stand the test of practicality. Once you have completed your study tasks and submitted your training results, the evaluation system will begin to quickly and accurately perform statistical assessments of your marks on the Databricks-Generative-AI-Engineer-Associate Exam Torrent so that you can arrange the learning tasks properly and focus on the targeted learning tasks with Databricks-Generative-AI-Engineer-Associate test questions.

**Valid Databricks-Generative-AI-Engineer-Associate Exam Discount:** <https://www.exam4labs.com/Databricks-Generative-AI-Engineer-Associate-practice-torrent.html>

- Databricks-Generative-AI-Engineer-Associate Actual Dumps  Databricks-Generative-AI-Engineer-Associate Exam Questions  Databricks-Generative-AI-Engineer-Associate VCE Exam Simulator  Search for  Databricks-Generative-AI-Engineer-Associate   and download exam materials for free through  [www.testkingpass.com](http://www.testkingpass.com)    Databricks-Generative-AI-Engineer-Associate New Test Materials
- 100% Pass Quiz 2026 Databricks Updated Databricks-Generative-AI-Engineer-Associate: Reliable Databricks Certified Generative AI Engineer Associate Exam Price  Search for  Databricks-Generative-AI-Engineer-Associate  and download it for free on  [www.pdfvce.com](http://www.pdfvce.com)  website  Databricks-Generative-AI-Engineer-Associate Test Simulator
- 100% Pass Quiz Databricks - Databricks-Generative-AI-Engineer-Associate - High Pass-Rate Reliable Databricks Certified Generative AI Engineer Associate Exam Price  Enter  [www.exam4labs.com](http://www.exam4labs.com)  and search for  Databricks-Generative-AI-Engineer-Associate  to download for free  Databricks-Generative-AI-Engineer-Associate Exam

## Questions

BTW, DOWNLOAD part of Exam4Labs Databricks-Generative-AI-Engineer-Associate dumps from Cloud Storage: [https://drive.google.com/open?id=1QAgv6nDBk5fumj\\_W1NDICY\\_MiEQwjaP](https://drive.google.com/open?id=1QAgv6nDBk5fumj_W1NDICY_MiEQwjaP)