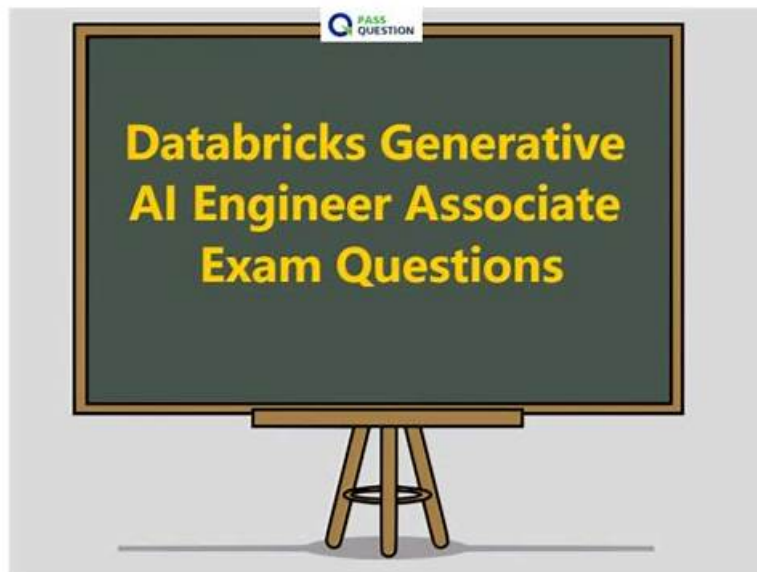


Relevant Databricks Databricks-Generative-AI-Engineer-Associate Questions & Practice Databricks-Generative-AI-Engineer-Associate Exam Online



P.S. Free 2025 Databricks Databricks-Generative-AI-Engineer-Associate dumps are available on Google Drive shared by BraindumpStudy: <https://drive.google.com/open?id=13n7IIIFZL4GsP9AYp19p9xgzvZcJPYUg>

The Databricks Certified Generative AI Engineer Associate prep torrent that we provide is compiled elaborately and highly efficient. You only need 20-30 hours to practice our Databricks-Generative-AI-Engineer-Associate exam torrent and then you can attend the exam. Among the people who prepare for the exam, many are office workers or the students. For the office worker, they are both busy in the job or their family; for the students, they possibly have to learn or do other things. But if they use our Databricks-Generative-AI-Engineer-Associate Test Prep, they won't need so much time to prepare the exam and master exam content in a short time. What they need to do is just to spare 1-2 hours to learn and practice every day and then pass the exam with Databricks-Generative-AI-Engineer-Associate test prep easily. It costs them little time and energy.

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

Topic	Details
Topic 1	• Application Development: In this topic, Generative AI Engineers learn about tools needed to extract data, Langchain • 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.
Topic 2	• 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.
Topic 3	• Governance: Generative AI Engineers who take the exam get knowledge about masking techniques, guardrail techniques, and legal • licensing requirements in this topic.
Topic 4	• 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.

2025 Realistic Relevant Databricks-Generative-AI-Engineer-Associate Questions - Practice Databricks Certified Generative AI Engineer Associate Exam Online Free PDF

If you start to prepare for the Databricks-Generative-AI-Engineer-Associate exam from books, then you will find that the content is too broad for you to cope with the exam questions. So, we just pick out the most important knowledge to learn. Through large numbers of practices, you will soon master the core knowledge of the Databricks-Generative-AI-Engineer-Associate Exam. It is important to review the questions you always choose mistakenly. You should concentrate on finishing all exercises once you are determined to pass the Databricks-Generative-AI-Engineer-Associate exam. And you will pass for sure as long as you study with our Databricks-Generative-AI-Engineer-Associate study guide carefully.

Databricks Certified Generative AI Engineer Associate Sample Questions (Q20-Q25):

NEW QUESTION # 20

After changing the response generating LLM in a RAG pipeline from GPT-4 to a model with a shorter context length that the company self-hosts, the Generative AI Engineer is getting the following error:

```
{"error_code": "BAD_REQUEST", "message": "Bad request: rpc error: code = InvalidArgument desc = prompt token count (4595) cannot exceed 4096..."}
```

What TWO solutions should the Generative AI Engineer implement without changing the response generating model? (Choose two.)

- A. Reduce the number of records retrieved from the vector database
- B. Use a smaller embedding model to generate
- C. Retrain the response generating model using ALiBi
- D. Reduce the maximum output tokens of the new model
- E. Decrease the chunk size of embedded documents

Answer: A,E

Explanation:

* Problem Context: After switching to a model with a shorter context length, the error message indicating that the prompt token count has exceeded the limit suggests that the input to the model is too large.

* Explanation of Options:

* Option A: Use a smaller embedding model to generate- This wouldn't necessarily address the issue of prompt size exceeding the model's token limit.

* Option B: Reduce the maximum output tokens of the new model- This option affects the output length, not the size of the input being too large.

* Option C: Decrease the chunk size of embedded documents- This would help reduce the size of each document chunk fed into the model, ensuring that the input remains within the model's context length limitations.

* Option D: Reduce the number of records retrieved from the vector database- By retrieving fewer records, the total input size to the model can be managed more effectively, keeping it within the allowable token limits.

* Option E: Retrain the response generating model using ALiBi- Retraining the model is contrary to the stipulation not to change the response generating model.

Options C and D are the most effective solutions to manage the model's shorter context length without changing the model itself, by adjusting the input size both in terms of individual document size and total documents retrieved.

NEW QUESTION # 21

A Generative AI Engineer would like an LLM to generate formatted JSON from emails. This will require parsing and extracting the following information: order ID, date, and sender email. Here's a sample email:

Date: April 23, 2024
Time: 4:22 PM
From: anjali.thayer@computex.org
To: cust_service@realtek.com
Subject: Shipment details

Hey there,

I have a shipment (order ID is CD34RFT) can you please send me an update?

Thank you,
Anjali

They will need to write a prompt that will extract the relevant information in JSON format with the highest level of output accuracy. Which prompt will do that?

- A. You will receive customer emails and need to extract date, sender email, and order ID. Return the extracted information in JSON format.
Here's an example: {"date": "April 16, 2024", "sender_email": "sarah.lee925@gmail.com", "order_id": "RE987D"}
- B. You will receive customer emails and need to extract date, sender email, and order ID. Return the extracted information in JSON format.
- C. You will receive customer emails and need to extract date, sender email, and order ID. You should return the date, sender email, and order ID information in JSON format.
- D. You will receive customer emails and need to extract date, sender email, and order ID. Return the extracted information in a human-readable format.

Answer: A

Explanation:

Problem Context: The goal is to parse emails to extract certain pieces of information and output this in a structured JSON format. Clarity and specificity in the prompt design will ensure higher accuracy in the LLM's responses.

Explanation of Options:

- * Option A: Provides a general guideline but lacks an example, which helps an LLM understand the exact format expected.
- * Option B: Includes a clear instruction and a specific example of the output format. Providing an example is crucial as it helps set the pattern and format in which the information should be structured, leading to more accurate results.
- * Option C: Does not specify that the output should be in JSON format, thus not meeting the requirement.
- * Option D: While it correctly asks for JSON format, it lacks an example that would guide the LLM on how to structure the JSON correctly.

Therefore, Option B is optimal as it not only specifies the required format but also illustrates it with an example, enhancing the likelihood of accurate extraction and formatting by the LLM.

NEW QUESTION # 22

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

Answer: D

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 # 23

A Generative AI Engineer has a provisioned throughput model serving endpoint as part of a RAG application and would like to monitor the serving endpoint's incoming requests and outgoing responses. The current approach is to include a micro-service in between the endpoint and the user interface to write logs to a remote server.

Which Databricks feature should they use instead which will perform the same task?

- **A. Inference Tables**
- B. DBSQL
- C. Lakeview
- D. Vector Search

Answer: A

Explanation:

Problem Context: The goal is to monitor the serving endpoint for incoming requests and outgoing responses in a provisioned throughput model serving endpoint within a Retrieval-Augmented Generation (RAG) application. The current approach involves using a microservice to log requests and responses to a remote server, but the Generative AI Engineer is looking for a more streamlined solution within Databricks.

Explanation of Options:

* Option A: Vector Search: This feature is used to perform similarity searches within vector databases.

It doesn't provide functionality for logging or monitoring requests and responses in a serving endpoint, so it's not applicable here.

* Option B: Lakeview: Lakeview is not a feature relevant to monitoring or logging request-response cycles for serving endpoints. It might be more related to viewing data in Databricks Lakehouse but doesn't fulfill the specific monitoring requirement.

* Option C: DBSQL: Databricks SQL (DBSQL) is used for running SQL queries on data stored in Databricks, primarily for analytics purposes. It doesn't provide the direct functionality needed to monitor requests and responses in real-time for an inference endpoint.

* Option D: Inference Tables: This is the correct answer. Inference Tables in Databricks are designed to store the results and metadata of inference runs. This allows the system to log incoming requests and outgoing responses directly within Databricks, making it an ideal choice for monitoring the behavior of a provisioned serving endpoint. Inference Tables can be queried and analyzed, enabling easier monitoring and debugging compared to a custom microservice.

Thus, Inference Tables are the optimal feature for monitoring request and response logs within the Databricks infrastructure for a model serving endpoint.

NEW QUESTION # 24

A Generative AI Engineer is creating an LLM-powered application that will need access to up-to-date news articles and stock prices.

The design requires the use of stock prices which are stored in Delta tables and finding the latest relevant news articles by searching the internet.

How should the Generative AI Engineer architect their LLM system?

- A. Download and store news articles and stock price information in a vector store. Use a RAG architecture to retrieve and generate at runtime.
- B. Use an LLM to summarize the latest news articles and lookup stock tickers from the summaries to find stock prices.
- **C. Create an agent with tools for SQL querying of Delta tables and web searching, provide retrieved values to an LLM for generation of response.**
- D. Query the Delta table for volatile stock prices and use an LLM to generate a search query to investigate potential causes of the stock volatility.

Answer: C

Explanation:

To build an LLM-powered system that accesses up-to-date news articles and stock prices, the best approach is to create an agent that has access to specific tools (option D).

* **Agent with SQL and Web Search Capabilities:** By using an agent-based architecture, the LLM can interact with external tools. The agent can query Delta tables (for up-to-date stock prices) via SQL and perform web searches to retrieve the latest news articles. This modular approach ensures the system can access both structured (stock prices) and unstructured (news) data sources dynamically.

* **Why This Approach Works:**

* **SQL Queries for Stock Prices:** Delta tables store stock prices, which the agent can query directly for the latest data.

* **Web Search for News:** For news articles, the agent can generate search queries and retrieve the most relevant and recent articles, then pass them to the LLM for processing.

* **Why Other Options Are Less Suitable:**

* **A (Summarizing News for Stock Prices):** This convoluted approach would not ensure accuracy when retrieving stock prices, which are already structured and stored in Delta tables.

* **B (Stock Price Volatility Queries):** While this could retrieve relevant information, it doesn't address how to obtain the most up-to-date news articles.

* **C (Vector Store):** Storing news articles and stock prices in a vector store might not capture the real-time nature of stock data and news updates, as it relies on pre-existing data rather than dynamic querying.

Thus, using an agent with access to both SQL for querying stock prices and web search for retrieving news articles is the best approach for ensuring up-to-date and accurate responses.

NEW QUESTION # 25

.....

Download the free Databricks-Generative-AI-Engineer-Associate pdf demo file of BraindumpStudy brain dumps. Checking the worth of the Databricks-Generative-AI-Engineer-Associate exam questions and learns the format of questions and answers. A few moments are enough to introduce you to the excellent of the Databricks-Generative-AI-Engineer-Associate Brain Dumps and the authenticity and relevance of the information contained in them.

Practice Databricks-Generative-AI-Engineer-Associate Exam Online: https://www.braindumpstudy.com/Databricks-Generative-AI-Engineer-Associate_braindumps.html

- Passing Databricks Certified Generative AI Engineer Associate actual test, valid Databricks-Generative-AI-Engineer-Associate test braindump □ Open □ www.exams4collection.com □ enter ✨ Databricks-Generative-AI-Engineer-Associate □ ✨ □ and obtain a free download □ Databricks-Generative-AI-Engineer-Associate Dumps Vce
- Databricks-Generative-AI-Engineer-Associate Dumps Vce □ Databricks-Generative-AI-Engineer-Associate Reliable Exam Tutorial □ Reliable Databricks-Generative-AI-Engineer-Associate Test Prep □ Go to website □ www.pdfvce.com □ open and search for “Databricks-Generative-AI-Engineer-Associate” to download for free □ □ Databricks-Generative-AI-Engineer-Associate Test Assessment
- Actual Databricks-Generative-AI-Engineer-Associate Test □ Databricks-Generative-AI-Engineer-Associate Free Braindumps □ Databricks-Generative-AI-Engineer-Associate Test Assessment □ Easily obtain free download of 「 Databricks-Generative-AI-Engineer-Associate 」 by searching on ➡ www.prep4pass.com □ □ Test Databricks-Generative-AI-Engineer-Associate Cram
- Free PDF Quiz 2025 Databricks-Generative-AI-Engineer-Associate: Databricks Certified Generative AI Engineer Associate Updated Relevant Questions □ Open ⇒ www.pdfvce.com ⇐ enter « Databricks-Generative-AI-Engineer-Associate » and obtain a free download □ Databricks-Generative-AI-Engineer-Associate Reliable Exam Tutorial
- Free PDF Quiz 2025 Databricks-Generative-AI-Engineer-Associate: Databricks Certified Generative AI Engineer Associate Updated Relevant Questions □ Search for ✨ Databricks-Generative-AI-Engineer-Associate □ ✨ □ and easily obtain a

free download on ⇒ www.prep4away.com ⇐ □ Databricks-Generative-AI-Engineer-Associate Test Dumps

- Databricks-Generative-AI-Engineer-Associate Free Braindumps □ Databricks-Generative-AI-Engineer-Associate Dumps Vce □ Databricks-Generative-AI-Engineer-Associate Prepaway Dumps □ Go to website 【 www.pdfvce.com 】 open and search for > Databricks-Generative-AI-Engineer-Associate □ to download for free ☞ Reliable Databricks-Generative-AI-Engineer-Associate Dumps Questions
- Databricks-Generative-AI-Engineer-Associate Valuable Feedback □ Actual Databricks-Generative-AI-Engineer-Associate Test □□ Databricks-Generative-AI-Engineer-Associate Reliable Exam Tutorial □ Immediately open > www.real4dumps.com ◁ and search for (Databricks-Generative-AI-Engineer-Associate) to obtain a free download □ □ Databricks-Generative-AI-Engineer-Associate Test Assessment
- Reliable Databricks-Generative-AI-Engineer-Associate Dumps Questions □ Test Databricks-Generative-AI-Engineer-Associate Cram □ Databricks-Generative-AI-Engineer-Associate Valuable Feedback □ Go to website 「 www.pdfvce.com 」 open and search for (Databricks-Generative-AI-Engineer-Associate) to download for free □ □ Databricks-Generative-AI-Engineer-Associate Prepaway Dumps
- Reliable Databricks-Generative-AI-Engineer-Associate Exam Simulations □ Databricks-Generative-AI-Engineer-Associate Latest Examprep □ Databricks-Generative-AI-Engineer-Associate Test Torrent □ Search for > Databricks-Generative-AI-Engineer-Associate ◁ and easily obtain a free download on □ www.prep4pass.com □ □ Databricks-Generative-AI-Engineer-Associate Test Dumps
- Databricks-Generative-AI-Engineer-Associate Free Braindumps □ Databricks-Generative-AI-Engineer-Associate Latest Braindumps Sheet □ Test Databricks-Generative-AI-Engineer-Associate Price □ Open [www.pdfvce.com] and search for [Databricks-Generative-AI-Engineer-Associate] to download exam materials for free □ Databricks-Generative-AI-Engineer-Associate Reliable Exam Tutorial
- High Pass Rate Databricks Certified Generative AI Engineer Associate Test Torrent is Convenient to Download - www.examcollectionpass.com □ Easily obtain ☼ Databricks-Generative-AI-Engineer-Associate □☼□ for free download through 「 www.examcollectionpass.com 」 □ Reliable Databricks-Generative-AI-Engineer-Associate Test Prep
- study.stcs.edu.np, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.xiaokedou21.com, www.stes.tyc.edu.tw, pct.edu.pk, Disposable vapes

P.S. Free 2025 Databricks Databricks-Generative-AI-Engineer-Associate dumps are available on Google Drive shared by BraindumpStudy: <https://drive.google.com/open?id=13n7IIIFZL4GsP9AYp19p9xgzvZcJPYUg>