

Valid Databricks-Generative-AI-Engineer-Associate Exam Camp - Databricks-Generative-AI-Engineer-Associate Latest Test Cost



After choose Actual4Exams's Databricks-Generative-AI-Engineer-Associate exam training materials, you can get the latest edition of Databricks-Generative-AI-Engineer-Associate exam dumps and answers. The accuracy rate of Actual4Exams Databricks-Generative-AI-Engineer-Associate exam training materials can ensure you to Pass Databricks-Generative-AI-Engineer-Associate Test. After you purchase our Databricks-Generative-AI-Engineer-Associate test training materials, if you fail Databricks-Generative-AI-Engineer-Associate exam certification or there are any quality problems of Databricks-Generative-AI-Engineer-Associate exam dumps, we guarantee a full refund.

The 24/7 support team is just an e-mail away for our customers so that they can contact us anytime. Our team will solve all of their issues as quickly as possible. Free demos and up to 1 year of free updates of our Databricks Exams are also available at Actual4Exams. Buy updated and Real Databricks-Generative-AI-Engineer-Associate Exam Questions now and earn your dream Databricks-Generative-AI-Engineer-Associate certification with Actual4Exams!

>> Valid Databricks-Generative-AI-Engineer-Associate Exam Camp <<

Databricks-Generative-AI-Engineer-Associate reliable training dumps & Databricks-Generative-AI-Engineer-Associate latest practice vce & Databricks-Generative-AI-Engineer-Associate valid study torrent

Our Databricks-Generative-AI-Engineer-Associate exam torrent is highly regarded in the market of this field and come with high recommendation. Choosing our Databricks-Generative-AI-Engineer-Associate exam guide will be a very promising start for you to begin your exam preparation because our Databricks-Generative-AI-Engineer-Associate practice materials with high reput. We remunerate exam candidates who fail the Databricks-Generative-AI-Engineer-Associate Exam Torrent after choosing our Databricks-Generative-AI-Engineer-Associate study tools, which kind of situation is rare but we still support your dream and help you avoid any kind of loss. Just try it do it, and we will be your strong backup.

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

Topic	Details
Topic 1	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 2	• Governance: Generative AI Engineers who take the exam get knowledge about masking techniques, guardrail techniques, and legal • licensing requirements in this topic.
Topic 3	• Assembling and Deploying Applications: In this topic, Generative AI Engineers get knowledge about coding a chain using a pyfunc mode, coding a simple chain using langchain, and coding a simple chain according to requirements. Additionally, the topic focuses on basic elements needed to create a RAG application. Lastly, the topic addresses sub-topics about registering the model to Unity Catalog using MLflow.
Topic 4	• 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 5	• 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.

Databricks Certified Generative AI Engineer Associate Sample Questions (Q55-Q60):

NEW QUESTION # 55

When developing an LLM application, it's crucial to ensure that the data used for training the model complies with licensing requirements to avoid legal risks.

Which action is NOT appropriate to avoid legal risks?

- A. Use any available data you personally created which is completely original and you can decide what license to use.
- B. Only use data explicitly labeled with an open license and ensure the license terms are followed.
- C. Reach out to the data curators directly after you have started using the trained model to let them know.
- D. Reach out to the data curators directly before you have started using the trained model to let them know.

Answer: C

Explanation:

* Problem Context: When using data to train a model, it's essential to ensure compliance with licensing to avoid legal risks. Legal issues can arise from using data without permission, especially when it comes from third-party sources.

* Explanation of Options:

* Option A: Reaching out to data curators before using the data is an appropriate action. This allows you to ensure you have permission or understand the licensing terms before starting to use the data in your model.

* Option B: Using original data that you personally created is always a safe option. Since you have full ownership over the data, there are no legal risks, as you control the licensing.

* Option C: Using data that is explicitly labeled with an open license and adhering to the license terms is a correct and recommended approach. This ensures compliance with legal requirements.

* Option D: Reaching out to the data curators after you have already started using the trained model is not appropriate. If you've already used the data without understanding its licensing terms, you may have already violated the terms of use, which could lead to legal complications. It's essential to clarify the licensing terms before using the data, not after.

Thus, Option D is not appropriate because it could expose you to legal risks by using the data without first obtaining the proper licensing permissions.

NEW QUESTION # 56

A Generative AI Engineer has already trained an LLM on Databricks and it is now ready to be deployed. Which of the following steps correctly outlines the easiest process for deploying a model on Databricks?

- **A. Log the model using MLflow during training, directly register the model to Unity Catalog using the MLflow API, and start a serving endpoint**
- B. Log the model as a pickle object, upload the object to Unity Catalog Volume, register it to Unity Catalog using MLflow, and start a serving endpoint
- C. Save the model along with its dependencies in a local directory, build the Docker image, and run the Docker container
- D. Wrap the LLM's prediction function into a Flask application and serve using Gunicorn

Answer: A

Explanation:

* Problem Context: The goal is to deploy a trained LLM on Databricks in the simplest and most integrated manner.

* Explanation of Options:

* Option A: This method involves unnecessary steps like logging the model as a pickle object, which is not the most efficient path in a Databricks environment.

* Option B: Logging the model with MLflow during training and then using MLflow's API to register and start serving the model is straightforward and leverages Databricks' built-in functionalities for seamless model deployment.

* Option C: Building and running a Docker container is a complex and less integrated approach within the Databricks ecosystem.

* Option D: Using Flask and Gunicorn is a more manual approach and less integrated compared to the native capabilities of Databricks and MLflow.

Option B provides the most straightforward and efficient process, utilizing Databricks' ecosystem to its full advantage for deploying models.

NEW QUESTION # 57

A small and cost-conscious startup in the cancer research field wants to build a RAG application using Foundation Model APIs. Which strategy would allow the startup to build a good-quality RAG application while being cost-conscious and able to cater to customer needs?

- **A. Pick a smaller LLM that is domain-specific**
- B. Use the largest LLM possible because that gives the best performance for any general queries
- C. Limit the number of queries a customer can send per day
- D. Limit the number of relevant documents available for the RAG application to retrieve from

Answer: A

Explanation:

For a small, cost-conscious startup in the cancer research field, choosing a domain-specific and smaller LLM is the most effective strategy. Here's why B is the best choice:

* Domain-specific performance: A smaller LLM that has been fine-tuned for the domain of cancer research will outperform a general-purpose LLM for specialized queries. This ensures high-quality responses without needing to rely on a large, expensive LLM.

* Cost-efficiency: Smaller models are cheaper to run, both in terms of compute resources and API usage costs. A domain-specific smaller LLM can deliver good quality responses without the need for the extensive computational power required by larger models.

* Focused knowledge: In a specialized field like cancer research, having an LLM tailored to the subject matter provides better relevance and accuracy for queries, while keeping costs low. Large, general-purpose LLMs may provide irrelevant information, leading to inefficiency and higher costs.

This approach allows the startup to balance quality, cost, and customer satisfaction effectively, making it the most suitable strategy.

NEW QUESTION # 58

A Generative AI Engineer is using an LLM to classify species of edible mushrooms based on text descriptions of certain features. The model is returning accurate responses in testing and the Generative AI Engineer is confident they have the correct list of possible labels, but the output frequently contains additional reasoning in the answer when the Generative AI Engineer only wants to return the label with no additional text.

Which action should they take to elicit the desired behavior from this LLM?

- **A. Use zero shot prompting to instruct the model on expected output format**
- B. Use zero shot chain-of-thought prompting to prevent a verbose output format

- C. Use a system prompt to instruct the model to be succinct in its answer
- D. Use few shot prompting to instruct the model on expected output format

Answer: C

Explanation:

The LLM classifies mushroom species accurately but includes unwanted reasoning text, and the engineer wants only the label. Let's assess how to control output format effectively.

- * Option A: Use few shot prompting to instruct the model on expected output format
- * Few-shot prompting provides examples (e.g., input: description, output: label). It can work but requires crafting multiple examples, which is effort-intensive and less direct than a clear instruction.
- * Databricks Reference: "Few-shot prompting guides LLMs via examples, effective for format control but requires careful design" ("Generative AI Cookbook").
- * Option B: Use zero shot prompting to instruct the model on expected output format
- * Zero-shot prompting relies on a single instruction (e.g., "Return only the label") without examples. It's simpler than few-shot but may not consistently enforce succinctness if the LLM's default behavior is verbose.
- * Databricks Reference: "Zero-shot prompting can specify output but may lack precision without examples" ("Building LLM Applications with Databricks").
- * Option C: Use zero shot chain-of-thought prompting to prevent a verbose output format
- * Chain-of-Thought (CoT) encourages step-by-step reasoning, which increases verbosity-opposite to the desired outcome. This contradicts the goal of label-only output.
- * Databricks Reference: "CoT prompting enhances reasoning but often results in detailed responses" ("Databricks Generative AI Engineer Guide").
- * Option D: Use a system prompt to instruct the model to be succinct in its answer
- * A system prompt (e.g., "Respond with only the species label, no additional text") sets a global instruction for the LLM's behavior. It's direct, reusable, and effective for controlling output style across queries.
- * Databricks Reference: "System prompts define LLM behavior consistently, ideal for enforcing concise outputs" ("Generative AI Cookbook," 2023).

Conclusion: Option D is the most effective and straightforward action, using a system prompt to enforce succinct, label-only responses, aligning with Databricks' best practices for output control.

NEW QUESTION # 59

A Generative AI Engineer has been asked to build an LLM-based question-answering application. The application should take into account new documents that are frequently published. The engineer wants to build this application with the least cost and least development effort and have it operate at the lowest cost possible.

Which combination of chaining components and configuration meets these requirements?

- A. The LLM needs to be frequently with the new documents in order to provide most up-to-date answers.
- B. For the question-answering application, prompt engineering and an LLM are required to generate answers.
- C. For the application a prompt, a retriever, and an LLM are required. The retriever output is inserted into the prompt which is given to the LLM to generate answers.
- D. For the application a prompt, an agent and a fine-tuned LLM are required. The agent is used by the LLM to retrieve relevant content that is inserted into the prompt which is given to the LLM to generate answers.

Answer: C

Explanation:

Problem Context: The task is to build an LLM-based question-answering application that integrates new documents frequently with minimal costs and development efforts.

Explanation of Options:

- * Option A: Utilizes a prompt and a retriever, with the retriever output being fed into the LLM. This setup is efficient because it dynamically updates the data pool via the retriever, allowing the LLM to provide up-to-date answers based on the latest documents without needing to frequently retrain the model. This method offers a balance of cost-effectiveness and functionality.
 - * Option B: Requires frequent retraining of the LLM, which is costly and labor-intensive.
 - * Option C: Only involves prompt engineering and an LLM, which may not adequately handle the requirement for incorporating new documents unless it's part of an ongoing retraining or updating mechanism, which would increase costs.
 - * Option D: Involves an agent and a fine-tuned LLM, which could be overkill and lead to higher development and operational costs.
- Option A is the most suitable as it provides a cost-effective, minimal development approach while ensuring the application remains up-to-date with new information.

• • • • •

Databricks-Generative-AI-Engineer-Associate Latest Test Cost: <https://www.actual4exams.com/Databricks-Generative-AI-Engineer-Associate-valid-dump.html>

- [illegible]

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, myportal.utt.edu.tt, elearning.cauqardho.edu.so, pct.edu.pk, yourstage.me, Disposable vapes