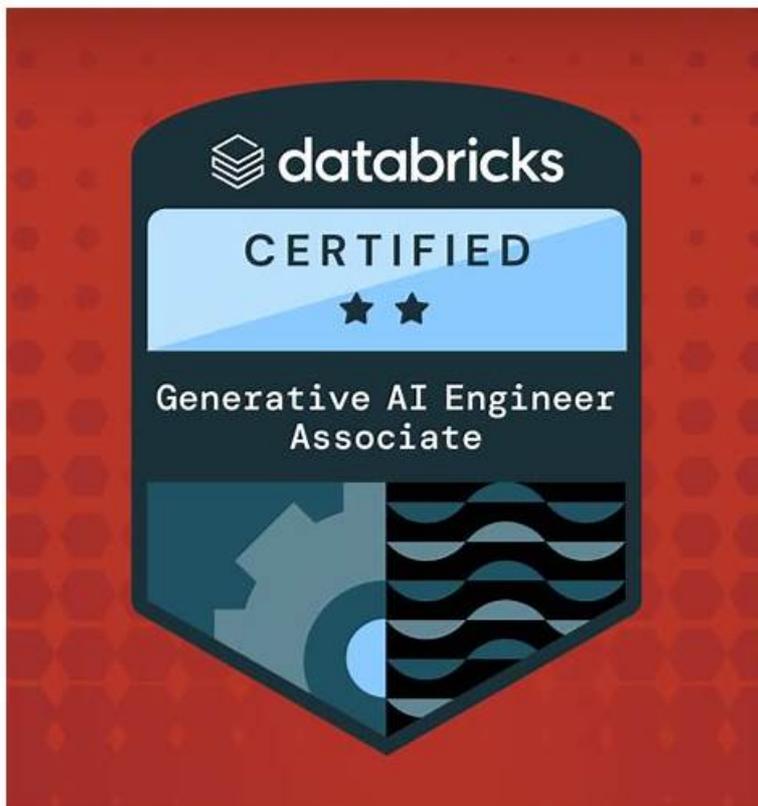


Databricks-Generative-AI-Engineer-Associate関連資格試験対応、Databricks-Generative-AI-Engineer-Associate日本語講座



P.S. TopexamがGoogle Driveで共有している無料かつ新しいDatabricks-Generative-AI-Engineer-Associateダウンロード: <https://drive.google.com/open?id=1k1JW6sdyYIatbK1uendXImaKmDe3NVMW>

お客様に自分に一番ふさわしいDatabricksのDatabricks-Generative-AI-Engineer-Associate試験の復習方式を提供するために、我々はDatabricksのDatabricks-Generative-AI-Engineer-Associateの資料の3つのバージョンを提供します。PDF、オンライン版とソフト版です。あなたの試験準備にヘルプを提供するのは常にあります。すべてのバージョンは無料のデモを提供します。そのほかに、どのバージョンでも全面的で最新版のDatabricksのDatabricks-Generative-AI-Engineer-Associateの資料を提供します。

Databricks Certified Generative AI Engineer Associate試験は大多数の受験者にとって難しい難題であることは広く受け入れられていますが、関連するDatabricks-Generative-AI-Engineer-Associate認定はこの分野の労働者にとって非常に重要であるため、多くの労働者はこの課題に取り組む必要があります。幸いなことに、この種の質問について心配する必要はありません。このWebサイトTopexamで最適なソリューションを見つけることができるので、Databricks-Generative-AI-Engineer-Associateトレーニング資料です。テクノロジー、人材、施設への継続的な投資により、当社Databricksの未来はこれまでになく輝かしく見えました。優れたDatabricks-Generative-AI-Engineer-Associate試験問題により、Databricks-Generative-AI-Engineer-Associate試験に合格します。

>> Databricks-Generative-AI-Engineer-Associate関連資格試験対応 <<

Databricks-Generative-AI-Engineer-Associate日本語講座 & Databricks-Generative-AI-Engineer-Associate専門知識訓練

古くから成功は準備のできる人のためにあると聞こえます。多くの人々は我々社のDatabricks-Generative-AI-Engineer-Associate問題集を介して、DatabricksのDatabricks-Generative-AI-Engineer-Associate試験資格認定を取得しました。しかも、この優位を持ってよい仕事を探しました。成功を受けたいあなたはすぐに行動しませんでしょうか? Databricks-Generative-AI-Engineer-Associate試験に興味があると、我々社Topexamをご覧になってください。

Databricks Databricks-Generative-AI-Engineer-Associate 認定試験の出題範囲:

トピック	出題範囲
トピック 1	<ul style="list-style-type: none"> • 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.
トピック 2	<ul style="list-style-type: none"> • 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.
トピック 3	<ul style="list-style-type: none"> • 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.
トピック 4	<ul style="list-style-type: none"> • 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 認定 Databricks-Generative-AI-Engineer-Associate 試験問題 (Q41-Q46):

質問 # 41

Generative AI Engineer at an electronics company just deployed a RAG application for customers to ask questions about products that the company carries. However, they received feedback that the RAG response often returns information about an irrelevant product.

What can the engineer do to improve the relevance of the RAG's response?

- A. Assess the quality of the retrieved context
- B. Use a different LLM to improve the generated response
- C. Implement caching for frequently asked questions
- D. Use a different semantic similarity search algorithm

正解: A

解説:

In a Retrieval-Augmented Generation (RAG) system, the key to providing relevant responses lies in the quality of the retrieved context. Here's why option A is the most appropriate solution:

Context Relevance:

The RAG model generates answers based on retrieved documents or context. If the retrieved information is about an irrelevant product, it suggests that the retrieval step is failing to select the right context. The Generative AI Engineer must first assess the quality of what is being retrieved and ensure it is pertinent to the query.

Vector Search and Embedding Similarity:

RAG typically uses vector search for retrieval, where embeddings of the query are matched against embeddings of product descriptions. Assessing the semantic similarity search process ensures that the closest matches are actually relevant to the query.

Fine-tuning the Retrieval Process:

By improving the retrieval quality, such as tuning the embeddings or adjusting the retrieval strategy, the system can return more accurate and relevant product information.

Why Other Options Are Less Suitable:

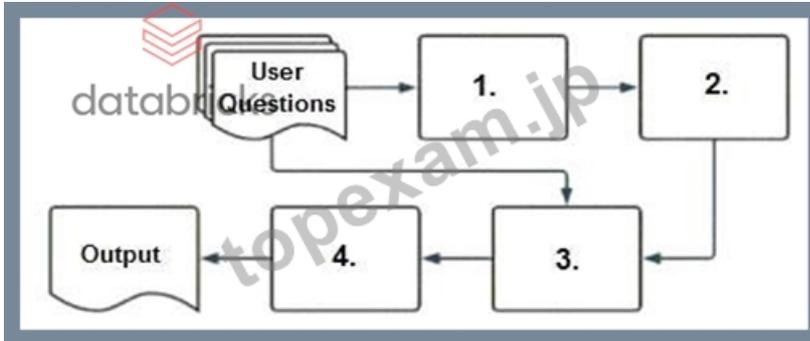
B (Caching FAQs): Caching can speed up responses for frequently asked questions but won't improve the relevance of the retrieved content for less frequent or new queries.

C (Use a Different LLM): Changing the LLM only affects the generation step, not the retrieval process, which is the core issue here.

D (Different Semantic Search Algorithm): This could help, but the first step is to evaluate the current retrieval context before replacing the search algorithm. Therefore, improving and assessing the quality of the retrieved context (option A) is the first step to fixing the issue of irrelevant product information.

質問 # 42

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



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

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

正解: B

解説:

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 misplace 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.

質問 # 43

A Generative AI Engineer is designing a chatbot for a gaming company that aims to engage users on its platform while its users play online video games.

Which metric would help them increase user engagement and retention for their platform?

- A. Randomness
- B. Lack of relevance
- C. Repetition of responses
- D. Diversity of responses

正解: D

解説:

In the context of designing a chatbot to engage users on a gaming platform, diversity of responses (option B) is a key metric to increase user engagement and retention. Here's why:

* Diverse and Engaging Interactions: A chatbot that provides varied and interesting responses will keep users engaged, especially in an interactive environment like a gaming platform. Gamers typically enjoy dynamic and evolving conversations, and diversity of responses helps prevent monotony, encouraging users to interact more frequently with the bot.

* Increasing Retention: By offering different types of responses to similar queries, the chatbot can create a sense of novelty and excitement, which enhances the user's experience and makes them more likely to return to the platform.

* Why Other Options Are Less Effective:

* A (Randomness): Random responses can be confusing or irrelevant, leading to frustration and reducing engagement.

* C (Lack of Relevance): If responses are not relevant to the user's queries, this will degrade the user experience and lead to disengagement.

* D (Repetition of Responses): Repetitive responses can quickly bore users, making the chatbot feel uninteresting and reducing the likelihood of continued interaction.

Thus, diversity of responses (option B) is the most effective way to keep users engaged and retain them on the platform.

質問 # 44

A Generative AI Engineer is designing a RAG application for answering user questions on technical regulations as they learn a new sport.

What are the steps needed to build this RAG application and deploy it?

- A. Ingest documents from a source -> Index the documents and save to Vector Search -> User submits queries against an LLM -> LLM retrieves relevant documents -> LLM generates a response -> Evaluate model -> Deploy it using Model Serving
- B. Ingest documents from a source -> Index the documents and saves to Vector Search -> User submits queries against an LLM -> LLM retrieves relevant documents -> Evaluate model -> LLM generates a response -> Deploy it using Model Serving
- C. User submits queries against an LLM -> Ingest documents from a source -> Index the documents and save to Vector Search -> LLM retrieves relevant documents -> LLM generates a response -> Evaluate model -> Deploy it using Model Serving
- D. Ingest documents from a source -> Index the documents and save to Vector Search -> Evaluate model -> Deploy it using Model Serving

正解: A

解説:

The Generative AI Engineer needs to follow a methodical pipeline to build and deploy a Retrieval-Augmented Generation (RAG) application. The steps outlined in option B accurately reflect this process:

Ingest documents from a source: This is the first step, where the engineer collects documents (e.g., technical regulations) that will be used for retrieval when the application answers user questions.

Index the documents and save to Vector Search: Once the documents are ingested, they need to be embedded using a technique like embeddings (e.g., with a pre-trained model like BERT) and stored in a vector database (such as Pinecone or FAISS). This enables fast retrieval based on user queries.

User submits queries against an LLM: Users interact with the application by submitting their queries. These queries will be passed to the LLM.

LLM retrieves relevant documents: The LLM works with the vector store to retrieve the most relevant documents based on their vector representations.

LLM generates a response: Using the retrieved documents, the LLM generates a response that is tailored to the user's question.

Evaluate model: After generating responses, the system must be evaluated to ensure the retrieved documents are relevant and the generated response is accurate. Metrics such as accuracy, relevance, and user satisfaction can be used for evaluation.

Deploy it using Model Serving: Once the RAG pipeline is ready and evaluated, it is deployed using a model-serving platform such as Databricks Model Serving. This enables real-time inference and response generation for users.

By following these steps, the Generative AI Engineer ensures that the RAG application is both efficient and effective for the task of answering technical regulation questions.

質問 # 45

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 Recall-Oriented-Understudy for Gisting Evaluation (ROUGE) scores of returned results for 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**

正解: D

解説:

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.

質問 # 46

.....

調査によると、当社の高く評価されているDatabricks-Generative-AI-Engineer-Associateテスト問題の成功は、簡単に操作できる練習システムへの尽力によるものです。候補者から寄せられたフィードバックのほとんどは、Databricks-Generative-AI-Engineer-Associateガイド急流が優れたプラクティスとシステムを実装し、より競争力のある新しい製品を発売する能力を強化していることを物語っています。Databricks-Generative-AI-Engineer-Associate試験ダンプに伴い、Q&Aはそれほど複雑ではありませんが、より重要な情報で受験者を教育します。これにより、Databricks-Generative-AI-Engineer-Associate試験に合格するための知識を深め、自己啓発を強化できます。

Databricks-Generative-AI-Engineer-Associate日本語講座: https://www.topexam.jp/Databricks-Generative-AI-Engineer-Associate_shiken.html

- 真実的なDatabricks-Generative-AI-Engineer-Associate関連資格試験対応一回合格-権威のあるDatabricks-Generative-AI-Engineer-Associate日本語講座 □ ➡ www.xhs1991.com □ を開き、⇒ Databricks-Generative-AI-Engineer-Associate ⇐ を入力して、無料でダウンロードしてくださいDatabricks-Generative-AI-Engineer-Associate リンクグローバル
- 真実的なDatabricks-Generative-AI-Engineer-Associate関連資格試験対応一回合格-権威のあるDatabricks-Generative-AI-Engineer-Associate日本語講座 □ ➡ www.goshiken.com □ サイトにて⇒ Databricks-Generative-AI-Engineer-Associate ⇐ 問題集を無料で使おうDatabricks-Generative-AI-Engineer-Associate日本語版問題解説
- 信賴的Databricks-Generative-AI-Engineer-Associate | 最高のDatabricks-Generative-AI-Engineer-Associate関連資格試験対応試験 | 試験の準備方法Databricks Certified Generative AI Engineer Associate日本語講座 □ 今すぐ □ www.goshiken.com □ で □ Databricks-Generative-AI-Engineer-Associate □ を検索して、無料でダウンロードして

