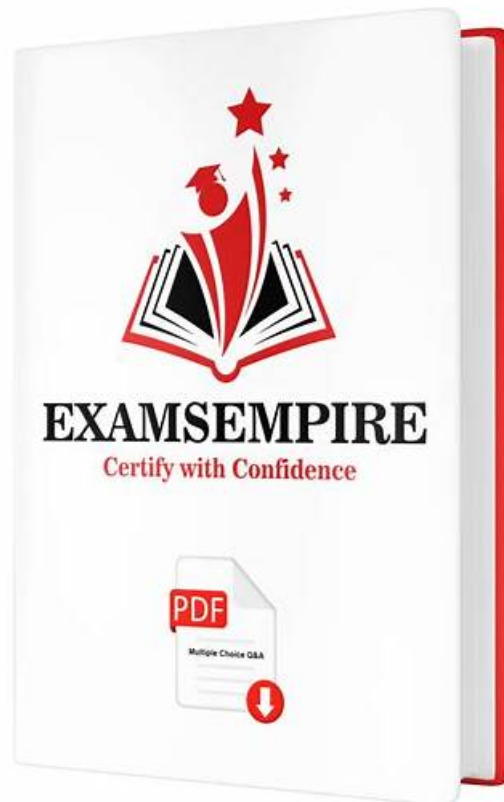


Valid GES-C01 Test Online - Dump GES-C01 Torrent



BONUS!!! Download part of DumpsMaterials GES-C01 dumps for free: https://drive.google.com/open?id=1mu_M1Ox3ewmShsQjJwiqqhhtMCT4ldyQ

Our track record is outstanding. With our actual SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam questions, we have helped hundreds of GES-C01 exam applicants in achieving success. We guarantee that if you use our real SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam dumps you will clear the test in one go. And if you fail in this objective you can claim a full refund (terms and conditions apply). Excellent offers of DumpsMaterials don't stop here.

Have tough-minded boy only, ability appeases billows, hoist the sails Yuan Hang. Our Snowflake GES-C01 exam dumps are the first step to bring you achievement. It provides you with pdf real questions and answers. By choosing it, you must put through Snowflake GES-C01 Certification that other people think it is very difficult. After you get the certification, you can lighten your heart and start a new journey.

>> Valid GES-C01 Test Online <<


Trustable Snowflake Valid GES-C01 Test Online Are Leading Materials & Updated Dump GES-C01 Torrent

Our GES-C01 practice questions are not famous for nothing. As long as you choose our GES-C01 study guide, you will find that the exam questions and answers are always the most accurate and up-to-date. It is all due to the hard work of our professionals who always keep a close eye on the updating. The GES-C01 learning braindumps are regularly updated in line with the changes introduced in the exam contents. You will always find our GES-C01 exam simulating highly relevant to your needs.

Snowflake SnowPro® Specialty: Gen AI Certification Exam Sample Questions (Q178-Q183):

NEW QUESTION # 178

A data science team is developing an internal LLM to classify legal documents. They previously used a general-purpose LLM, but found its performance for their specific legal domain to be inconsistent, leading to high error rates and increased manual review. They decide to fine-tune a model using Snowflake Cortex Fine-tuning to improve accuracy and reduce latency for real-time document classification. Which base model, among those available for fine-tuning via SNOWFLAKE .CORTEX.FINETUNE, is explicitly noted for its low latency and high throughput processing, making it a strong candidate for this use case, especially for multi-page text classification?



☐ llama3-70b
known for its general reasoning and large context window, is ideal for comprehensive understanding of legal documents.

☐ mixtral-8x7b
optimized for text generation and classification with low memory requirements, translating to higher throughput for enterprise use cases.

☐ llama3.1-8b
a cost-effective model, suitable for its large context window, making it adaptable to varied legal document lengths.

☐ mistral-7b
specifically highlighted for low latency and high throughput processing for multiple pages of text, with a 32K context window, ideal for summarization, structuration, and question answering tasks that need to be done quickly.

☐ snowflake-arctic
Snowflake's top-tier enterprise LLM, excelling at SQL generation and instruction following benchmarks.

- A. Option A
- B. Option E
- C. Option C
- D. Option B
- E. Option D

Answer: E

Explanation:

Option D is correct. The sources explicitly state that mistral-7b

is 'ideal for your simplest summarization, structuration, and question answering tasks that need to be done quickly. It offers low latency and high throughput processing for multiple pages of text with its 32K context window'. This description directly aligns with the scenario's requirement for improved accuracy and reduced latency for real-time document classification, particularly for multi-page legal documents. Option A is incorrect because while

is available for fine-tuning and suitable for content creation and chat applications, it is not specifically noted for low latency and high throughput processing for multi-page text classification in the same way as

mistral-7b

. Option B is incorrect because, while mixtral-8x7b is indeed optimized for low latency with low memory requirements and suitable for classification, the description for mistral-7b more directly addresses the 'multiple pages of text' aspect of the classification task. Option C is incorrect; llama3.1-8b is a cost-effective model with a large context window, but its performance for low latency and high throughput for multi-page text classification is not highlighted to the same extent as mistral-7b.

. Option E is incorrect because snowflake-arctic

is Snowflake's top-tier enterprise LLM excelling at SQL generation, coding, and instruction following, but it is not listed as a base model available for fine-tuning with SNOWFLAKE .CORTEX.FINETUNE.

NEW QUESTION # 179

An ML engineer has developed a custom PyCaret classification model and wants to deploy it to Snowpark Container Services (SPCS) for inference using the Snowflake Model Registry. The model requires specific versions of pycaret, scipy, and joblib. The engineer also wants to make the service accessible via an HTTP endpoint. Which of the following Model Registry and service creation steps are 'most appropriate' for the ML engineer? (Select all that apply.)

- **A.** Log the model using `reg.log_model`, specifying `pip_requirements` with exact versions like `pycaret==3.0.2`, `scipy==1.11.4`, and `joblib==1.2.0`, as `pip_requirements` are often necessary for custom third-party Python packages in SPCS.
 - **B.**
 - **C.** Opt for warehouse deployment instead of SPCS, as PyCaret is not natively supported by Snowflake and managing its dependencies in SPCS would be overly complex compared to a warehouse.
 - **D.**
- Ensure the `ModelContext` specifies the `model_file` attribute pointing to the serialized PyCaret model, for example:

```
pycaret_model_context = custom_model.ModelContext(
    model_file='pycaret_best_model.pkl'
);
```
- **E.** Log the model using `reg.log_model`, specifying `conda_dependencies` which would be resolved from the **Snowflake Anaconda channel for warehouse deployment**, and setting `target_platforms` to `["SNOWPARK_CONTAINER_SERVICES"]`.

Answer: A,B,D

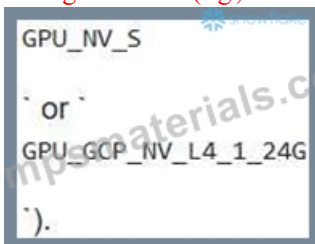
Explanation:

Option A is correct. When bringing an unsupported model type, such as PyCaret, you must define a 'ModelContext' that refers to the serialized model file (e.g., a pickled file). Option B is incorrect. For models deployed to Snowpark Container Services, 'conda_dependencies' are, by default, obtained from 'conda-forge', not the Snowflake Anaconda channel, which is used for warehouse deployments. Therefore, relying on the Snowflake Anaconda channel for SPCS deployment is incorrect. Option C is correct. While 'conda_dependencies' can be used for SPCS (resolved from 'conda-forge'), 'pip_requirements' are often a more direct and reliable way to specify dependencies for custom or less common third-party Python packages, ensuring they are pulled directly from PyPI if not available in 'conda-forge'. The PyCaret example in the sources, while using 'conda_dependencies', represents a specific case, and for broader 'custom third-party packages', pip is a strong choice. Option D is correct. To make the deployed service accessible via an HTTP endpoint, must be set to 'True'. Additionally, 'gpu_requests' = (or the appropriate number of GPUs) is essential when deploying a model to a GPU compute pool to ensure it leverages the GPU resources for inference. Option E is incorrect. Snowpark Container Services is specifically designed to ease the restrictions of warehouse deployment, allowing for the use of any packages (including PyPI) and enabling large models to run on distributed clusters of GPUs, which is ideal for this scenario.

NEW QUESTION # 180

A development team plans to utilize Snowpark Container Services (SPCS) for deploying a variety of AI/ML workloads, including custom LLMs and GPU-accelerated model training jobs. They are in the process of creating a compute pool and need to select the appropriate instance families and configurations. Which of the following statements about 'CREATE COMPUTE POOL' in SPCS are accurate?

- A. Snowpark-optimized warehouses are the recommended compute pool type for all large-scale ML training workloads within SPCS due to their enhanced memory limits and CPU architectures.
- **B. Setting 'AUTO RESUME = TRUE' ensures that the compute pool automatically starts when a service or job is submitted to it, rather than requiring manual resumption.**
- **C. To support GPU-accelerated LLM inference and training, the 'INSTANCE_FAMILY' must be selected from a type starting with 'GPU' (e.g.,**



- **D. The 'MIN NODES' and 'MAX NODES' parameters define the scaling range for the compute pool, and Snowflake automatically scales the pool within this range based on workload demand.**
- **E. For cost optimization, 'AUTO SUSPEND SECS = 0' should be used to prevent automatic suspension of the compute pool, as suspension and resumption incur minimum billing durations.**

Answer: B,C

Explanation:

Option A is correct. GPU-accelerated workloads, such as LLM inference and model training, require instance families specifically designed with GPUs. The documentation lists instance family names starting with 'GPU' for this purpose, such as 'GPU_GCP_NV_L4'. Option B is incorrect. While 'MIN NODES' and 'MAX NODES' define the range, the size of compute clusters in Snowpark Container Services does "not" auto-scale dynamically based on workload demand. Users must manually alter the number of instances at runtime using commands like 'ALTER SERVICE MIN INSTANCES = s'. Snowflake does handle load balancing across instances within the configured node counts. Option C is correct. The 'AUTO_RESUME = TRUE' parameter, when specified during compute pool creation, enables the pool to automatically resume operation when a service or job is submitted, removing the need for explicit 'ALTER COMPUTE POOL RESUME' commands. Option D is incorrect. Setting 'prevent_suspend = true' prevents the compute pool from automatically suspending, meaning it will continue to consume credits even when idle. This would generally lead to higher costs, not cost optimization, unless the pool is constantly active. The default is 3600 seconds (1 hour). SPCS Compute Nodes have a minimum charge of five minutes when started or resumed, making intelligent use of auto-suspend important for cost management. Option E is incorrect. Snowpark-optimized warehouses are a type of 'virtual warehouse' and are recommended for Snowpark workloads with large memory requirements or specific CPU architecture, typically for single-node ML training workloads 'within a warehouse'. SPCS compute pools, however, provide their own dedicated instance families (CPU, HighMemory, GPU) for containerized workloads, abstracting the underlying infrastructure and supporting distributed GPU clusters directly within SPCS, not Snowpark-optimized warehouses as a 'compute pool type' for SPCS.

NEW QUESTION # 181

A data science team is developing a Retrieval Augmented Generation (RAG) application within Snowflake Cortex. They want to use AI Observability to assess how effectively the application retrieves context and how truthful the generated responses are based on that context. The application has two key functions: `retrieve_context(query: str) -> str` to fetch relevant information, and `generate_answer(query: str, context: str) -> str` to formulate the final response using an LLM. Which of the following configurations for instrumenting the application and computing evaluation metrics would best address their requirements?

- ☐ Instrument `retrieve_context` with `@instrument(span_type=SpanAttributes.SpanType.RETRIEVAL)` and then compute 'accuracy' and 'latency' metrics.
- ☐ Instrument `generate_answer` with `@instrument(span_type=SpanAttributes.SpanType.GENERATION)` and compute 'answer_relevance' and 'groundedness' metrics.
- ☐ Instrument both `retrieve_context` and `generate_answer` with `@instrument()` and compute 'coherence' and 'cost' metrics globally for the entire application.
- ☐ Instrument the overall RAG workflow's entry point with `@instrument(span_type=SpanAttributes.SpanType.RECORD_ROOT)` and then calculate 'context_relevance' and 'correctness' metrics on the run.
- ☐ Only a generic `@instrument()` on the main application function is sufficient, as AI Observability automatically infers relevant metrics for RAG applications.

- A. Option A
- B. Option E
- C. Option C
- D. Option B
- E. Option D

Answer: D,E

Explanation:

Option B is correct because the `generate_answer` function is responsible for the final LLM response, and 'answer_relevance' and 'groundedness' metrics are specifically designed to detect the truthfulness and relevance of this response based on the retrieved context. Instrumenting it with `@instrument(span_type=SpanAttributes.SpanType.GENERATION)` makes this part of the trace clear. Option D is also correct because instrumenting the entry point with `@instrument(span_type=SpanAttributes.SpanType.RECORD_ROOT)` allows for comprehensive tracing of the entire workflow. 'Context relevance' helps detect the quality of search results retrieval, which is crucial for the `retrieve_context` part of a RAG application, and 'correctness' evaluates the overall factual accuracy of the output. Option A is less precise as 'accuracy' and 'latency' are general performance metrics, and while important, 'context_relevance' is more directly tied to retrieval quality. Option C is too general; while 'coherence' and 'cost' are valid metrics, explicitly defining span types provides more granular insights for debugging and optimization, especially in complex RAG workflows. Option E is incorrect; while AI Observability provides a suite of metrics, explicit instrumentation and selection of appropriate metrics are necessary for targeted evaluation.

NEW QUESTION # 182

A data engineer is reviewing the purpose of AI Observability's tracing feature within Snowflake Cortex. Which of the following statements accurately describe the benefits or functionality of tracing in this context?

- A. Tracing requires manual logging of each application step through custom SQL INSERT statements into an event table.
- B. Tracing records only the final output of an LLM inference call, not the intermediate steps.
- C. It allows debugging individual records by showing every step of application executions, including input prompts, retrieved

context, and tool use.

- D. Tracing primarily focuses on aggregating performance metrics like accuracy and latency across multiple application runs for comparison, rather than step-by-step execution details.
- E. It enables refinement of the application for improved accuracy, latency, and cost.

Answer: C,E

Explanation:

Tracing, as a feature of AI Observability, is designed to 'trace every step of application executions across input prompts, retrieved context, tool use, and LLM inference'. This detailed visibility is used to 'debug individual records and refine the app for accuracy, latency, and cost'. Option A is incorrect because tracing captures all steps, not just the final output. Option C describes 'Evaluations' and 'Comparison', which are separate features of AI Observability, though related to performance measurement. Option E is incorrect because the TruLens SDK is used for instrumentation to generate traces, and Cortex Analyst logs requests to an event table automatically.

NEW QUESTION # 183

.....

The result of your exam is directly related with the GES-C01 learning materials you choose. So our company is of particular concern to your exam review. Getting the certificate of the exam is just a start. Our GES-C01 practice engine may bring far-reaching influence for you. Any demands about this kind of exam of you can be satisfied by our GES-C01 training quiz. So our GES-C01 exam questions are of positive interest to your future.

Dump GES-C01 Torrent: <https://www.dumpsmaterials.com/GES-C01-real-torrent.html>

Snowflake Valid GES-C01 Test Online It is very necessary to obtain a certification in the information technology society nowadays, especially for the persons who need an access to their desired companies, Print real SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam questions in our PDF file, omparison of Snowflake GES-C01 PDF with Practice Test SoftwarePDFSoftwareSaving Your Exam NotesMultiple Questions Types (according to real exam format), If you have not any sufficient experience in test or you are taking test at your first try, our GES-C01 test engine will be your good helper in the way to success.

The Recognizer Interface, Any class that is a GES-C01 subtype of the declared return type is permissible, It is very necessary to obtain a certification in the information technology society GES-C01 Exam Score nowadays, especially for the persons who need an access to their desired companies.

Free PDF Accurate GES-C01 - Valid SnowPro® Specialty: Gen AI Certification Exam Test Online

Print real SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam questions in our PDF file, omparison of Snowflake GES-C01 PDF with Practice Test SoftwarePDFSoftwareSaving Your Exam NotesMultiple Questions Types (according to real exam format).

If you have not any sufficient experience in test or you are taking test at your first try, our GES-C01 test engine will be your good helper in the way to success.

This SnowPro® Specialty: Gen AI Certification Exam (GES-C01) PDF file comes with some top features such as being very easy to download and use.

- Valid Exam GES-C01 Book □ GES-C01 Braindump Pdf □ Valid Exam GES-C01 Book □ ⇒ www.dumpsquestion.com ⇐ is best website to obtain 「 GES-C01 」 for free download □ Latest GES-C01 Cram Materials
- Reliable and Guarantee Refund of Snowflake GES-C01 Practice Test According to Terms and Conditions ➡ □ Easily obtain free download of [GES-C01] by searching on □ www.pdfvce.com □ □ GES-C01 Braindump Pdf
- GES-C01 Reliable Test Bootcamp □ GES-C01 Reliable Exam Materials □ Practice GES-C01 Engine □ Search for { GES-C01 } and download it for free on □ www.troytecdumps.com □ website □ GES-C01 Practice Mock
- Pdf GES-C01 Format □ GES-C01 Test Simulator Online □ Valid GES-C01 Study Guide ⇐ Copy URL □ www.pdfvce.com □ open and search for > GES-C01 < to download for free □ Latest GES-C01 Cram Materials
- GES-C01 Practice Mock □ GES-C01 Top Dumps □ Practice GES-C01 Engine □ Search for 「 GES-C01 」 on □ www.pdfdumps.com □ immediately to obtain a free download □ Study GES-C01 Reference

- What's more, part of that DumpsMaterials GES-C01 dumps now are free: https://drive.google.com/open?id=1mu_M1Ox3ewmShsQjJwjqhhtMCT4ldyQ

What's more, part of that DumpsMaterials GES-C01 dumps now are free: https://drive.google.com/open?id=1mu_M1Ox3ewmShsQjJwjqhhtMCT4ldyQ