

Correct NCA-GENL Practice Questions & Marvelous NCA-GENL Latest Braindumps Files & Precise NVIDIA NVIDIA Generative AI LLMs



P.S. Free 2026 NVIDIA NCA-GENL dumps are available on Google Drive shared by TestkingPass:
https://drive.google.com/open?id=1HBjQw0uNfOzm0WvTAjz_qN72b3qRa0E

Once you have practiced on our NVIDIA Generative AI LLMs test questions, the system will automatically memorize and analyze all your practice. You must finish the model test in limited time. There have a timer on the right of the interface. Once you begin to do the exercises of the NCA-GENL test guide, the timer will start to work and count down. If you don't finish doing the exercises, all your exercises of the NCA-GENL Exam Questions will be delivered automatically. Then the system will generate a report according to your performance. You will clearly know where you are good at or not. Then you can make your own learning plans based on the report of the NCA-GENL test guide. Also, you will do more practices that you are not good at until you completely have no problem.

As you may know that the windows software of the NCA-GENL study materials only supports windows operating system. Also, it needs to run on Java environment. If the computer doesn't install JAVA, it will automatically download to ensure the normal running of the NCA-GENL Study Materials. What's more, all computers you have installed our study materials can run normally. Our NCA-GENL exam guide are cost-effective.

>> NCA-GENL Practice Questions <<

NCA-GENL Latest Braindumps Files, Reliable NCA-GENL Exam Cost

If you feel that you just don't have enough competitiveness to find a desirable job. Then it is time to strengthen your skills. Our NCA-GENL exam simulating will help you master the most popular skills in the job market. Then you will have a greater chance to find a desirable job. Also, it doesn't matter whether have basic knowledge about the NCA-GENL training quiz for the content of our NCA-GENL study guide contains all the exam keypoints which you need to cope with the real exam.

NVIDIA Generative AI LLMs Sample Questions (Q58-Q63):

NEW QUESTION # 58

What is the main difference between forward diffusion and reverse diffusion in diffusion models of Generative AI?

- A. Forward diffusion focuses on progressively injecting noise into data, while reverse diffusion focuses on generating new samples from the given noise vectors.
- B. Forward diffusion uses bottom-up processing, while reverse diffusion uses top-down processing to generate samples from noise vectors.
- C. Forward diffusion uses feed-forward networks, while reverse diffusion uses recurrent networks.
- D. Forward diffusion focuses on generating a sample from a given noise vector, while reverse diffusion reverses the process by estimating the latent space representation of a given sample.

Answer: A

Explanation:

Diffusion models, a class of generative AI models, operate in two phases: forward diffusion and reverse diffusion. According to NVIDIA's documentation on generative AI (e.g., in the context of NVIDIA's work on generative models), forward diffusion progressively injects noise into a data sample (e.g., an image or text embedding) over multiple steps, transforming it into a noise distribution. Reverse diffusion, conversely, starts with a noise vector and iteratively denoises it to generate a new sample that resembles the training data distribution. This process is central to models like DDPM (Denoising Diffusion Probabilistic Models). Option A is incorrect, as forward diffusion adds noise, not generates samples. Option B is false, as diffusion models typically use convolutional or transformer-based architectures, not recurrent networks. Option C is misleading, as diffusion does not align with bottom-up/top-down processing paradigms.

References:

NVIDIA Generative AI Documentation: <https://www.nvidia.com/en-us/ai-data-science/generative-ai/> Ho, J., et al. (2020). "Denoising Diffusion Probabilistic Models."

NEW QUESTION # 59

What are the main advantages of instructed large language models over traditional, small language models (< 300M parameters)? (Pick the 2 correct responses)

- A. Trained without the need for labeled data.
- B. It is easier to explain the predictions.
- C. Cheaper computational costs during inference.
- D. Smaller latency, higher throughput.
- E. Single generic model can do more than one task.

Answer: C,E

Explanation:

Instructed large language models (LLMs), such as those supported by NVIDIA's NeMo framework, have significant advantages over smaller, traditional models:

* Option D: LLMs often have cheaper computational costs during inference for certain tasks because they can generalize across multiple tasks without requiring task-specific retraining, unlike smaller models that may need separate models per task.

References:

NVIDIA NeMo Documentation: <https://docs.nvidia.com/deeplearning/nemo/user-guide/docs/en/stable/nlp/intro.html>

Brown, T., et al. (2020). "Language Models are Few-Shot Learners."

NEW QUESTION # 60

Which of the following claims is correct about TensorRT and ONNX?

- A. TensorRT is used for model creation and ONNX is used for model deployment.
- B. TensorRT is used for model deployment and ONNX is used for model interchange.
- C. TensorRT is used for model deployment and ONNX is used for model creation.
- D. TensorRT is used for model creation and ONNX is used for model interchange.

Answer: B

Explanation:

NVIDIA TensorRT is a deep learning inference library used to optimize and deploy models for high-performance inference, while ONNX (Open Neural Network Exchange) is a format for model interchange, enabling models to be shared across different frameworks, as covered in NVIDIA's Generative AI and LLMs course. TensorRT optimizes models (e.g., via layer fusion and quantization) for deployment on NVIDIA GPUs, while ONNX ensures portability by providing a standardized model representation. Option B is incorrect, as ONNX is not used for model creation but for interchange. Option C is wrong, as TensorRT is not for model creation but optimization and deployment. Option D is inaccurate, as ONNX is not for deployment but for model sharing. The course notes: "TensorRT optimizes and deploys deep learning models for inference, while ONNX enables model interchange across frameworks for portability." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

NEW QUESTION # 61

In the context of language models, what does an autoregressive model predict?

- A. The probability of the next token using a Monte Carlo sampling of past tokens.
- B. The next token solely using recurrent network or LSTM cells.
- C. The probability of the next token by looking at the previous and future input tokens.
- **D. The probability of the next token in a text given the previous tokens.**

Answer: D

Explanation:

Autoregressive models are a cornerstone of modern language modeling, particularly in large language models (LLMs) like those discussed in NVIDIA's Generative AI and LLMs course. These models predict the probability of the next token in a sequence based solely on the preceding tokens, making them inherently sequential and unidirectional. This process is often referred to as "next-token prediction," where the model learns to generate text by estimating the conditional probability distribution of the next token given the context of all previous tokens. For example, given the sequence "The cat is," the model predicts the likelihood of the next word being "on," "in," or another token. This approach is fundamental to models like GPT, which rely on autoregressive decoding to generate coherent text. Unlike bidirectional models (e.g., BERT), which consider both previous and future tokens, autoregressive models focus only on past tokens, making option D incorrect. Options B and C are also inaccurate, as Monte Carlo sampling is not a standard method for next-token prediction in autoregressive models, and the prediction is not limited to recurrent networks or LSTM cells, as modern LLMs often use Transformer architectures. The course emphasizes this concept in the context of Transformer-based NLP: "Learn the basic concepts behind autoregressive generative models, including next-token prediction and its implementation within Transformer-based models." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

NEW QUESTION # 62

What is confidential computing?

- A. A method for interpreting and integrating various forms of data in AI systems.
- **B. A technique for securing computer hardware and software from potential threats.**
- C. A process for designing and applying AI systems in a manner that is explainable, fair, and verifiable.
- D. A technique for aligning the output of the AI models with human beliefs.

Answer: B

Explanation:

Confidential computing is a technique for securing computer hardware and software from potential threats by protecting data in use, as covered in NVIDIA's Generative AI and LLMs course. It ensures that sensitive data, such as model weights or user inputs, remains encrypted during processing, using technologies like secure enclaves or trusted execution environments (e.g., NVIDIA H100 GPUs with confidential computing capabilities). This enhances the security of AI systems. Option B is incorrect, as it describes Trustworthy AI principles, not confidential computing. Option C is wrong, as aligning outputs with human beliefs is unrelated to security. Option D is inaccurate, as data integration is not the focus of confidential computing. The course notes: "Confidential computing secures AI systems by protecting data in use, leveraging trusted execution environments to safeguard sensitive information during processing." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

NEW QUESTION # 63

.....

TestkingPass NCA-GENL exam training materials can help you save a lot of time and energy, and make you yield twice the result with half effort to pass NCA-GENL certification exam. After you purchase our NCA-GENL exam dumps, we will also provide one year free renewal service for you. If there's any quality problem in NCA-GENL Exam Dumps you buy or you fail NCA-GENL certification exam, we promise to give you a full refund unconditionally.

NCA-GENL Latest Braindumps Files: <https://www.testkingpass.com/NCA-GENL-testking-dumps.html>

We are ready for providing the best NCA-GENL test guide materials for you, NVIDIA NCA-GENL Practice Questions You can print the PDF version out, NVIDIA NCA-GENL Practice Questions =It is acknowledged that high-quality service after sales plays a vital role in enhancing the relationship between the company and customers, NVIDIA NCA-GENL Practice Questions We aim to serve every customer heart and soul.

