

# NCA-GENL Trustworthy Exam Torrent | Test NCA-GENL Assessment

## Video 2: About NCA-GENL Certification Key Features



DOWNLOAD the newest Exams4Collection NCA-GENL PDF dumps from Cloud Storage for free:  
<https://drive.google.com/open?id=1b0flGCbm4jsEWFBqdbOwa8tOsHJZTlpH>

With so many methods can boost individual competitiveness, people may be confused, which can really bring them a glamorous work or brighter future? We are here to tell you that a NCA-GENL certification definitely has everything to gain and nothing to lose for everyone. You might have seen lots of advertisements about NCA-GENL learning question, there are so many types of NCA-GENL exam material in the market, why you should choose us? Our reasons are as follow. Our NCA-GENL test guide is test-oriented, which makes the preparation become highly efficient.

Most of the experts in our company have been studying in the professional field for many years and have accumulated much experience in our NCA-GENL practice questions. Our company is considerably cautious in the selection of talent and always hires employees with store of specialized knowledge and skills. All the members of our experts and working staff maintain a high sense of responsibility, which is why there are so many people choose our NCA-GENL Exam Materials and to be our long-term partner.

>> NCA-GENL Trustworthy Exam Torrent <<

## 100% Pass NVIDIA - Unparalleled NCA-GENL - NVIDIA Generative AI LLMs Trustworthy Exam Torrent

Before you take the NCA-GENL exam, you only need to spend 20 to 30 hours to practice, so you can schedule time to balance learning and other things. Of course, you care more about your passing rate. If you choose our NCA-GENL exam guide, under the guidance of our NCA-GENL exam torrent, we have the confidence to guarantee a passing rate of over 99%. Our NCA-GENL Quiz prep is compiled by experts based on the latest changes in the teaching syllabus and theories and practices. So our NCA-GENL quiz prep is quality-assured, focused, and has a high hit rate.

## NVIDIA Generative AI LLMs Sample Questions (Q15-Q20):

### NEW QUESTION # 15

Why do we need positional encoding in transformer-based models?

- A. To prevent overfitting of the model.
- **B. To represent the order of elements in a sequence.**
- C. To reduce the dimensionality of the input data.
- D. To increase the throughput of the model.

**Answer: B**

Explanation:

Positional encoding is a critical component in transformer-based models because, unlike recurrent neural networks (RNNs), transformers process input sequences in parallel and lack an inherent sense of word order.

Positional encoding addresses this by embedding information about the position of each token in the sequence, enabling the model to understand the sequential relationships between tokens. According to the original transformer paper ("Attention is All You Need" by Vaswani et al., 2017), positional encodings are added to the input embeddings to provide the model with information about the relative or absolute position of tokens. NVIDIA's documentation on transformer-based models, such as those supported by the NeMo framework, emphasizes that positional encodings are typically implemented using sinusoidal functions or learned embeddings to preserve sequence order, which is essential for tasks like natural language processing (NLP). Options B, C, and D are incorrect because positional encoding does not address overfitting, dimensionality reduction, or throughput directly; these are handled by other techniques like regularization, dimensionality reduction methods, or hardware optimization.

References:

Vaswani, A., et al. (2017). "Attention is All You Need."

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

### NEW QUESTION # 16

"Hallucinations" is a term coined to describe when LLM models produce what?

- A. Grammatically incorrect or broken outputs.
- B. Images from a prompt description.
- C. Correct sounding results that are wrong.
- D. Outputs are only similar to the input data.

**Answer: C**

Explanation:

In the context of LLMs, "hallucinations" refer to outputs that sound plausible and correct but are factually incorrect or fabricated, as emphasized in NVIDIA's Generative AI and LLMs course. This occurs when models generate responses based on patterns in training data without grounding in factual knowledge, leading to misleading or invented information. Option A is incorrect, as hallucinations are not about similarity to input data but about factual inaccuracies. Option B is wrong, as hallucinations typically refer to text, not image generation. Option D is inaccurate, as hallucinations are grammatically coherent but factually wrong. The course states: "Hallucinations in LLMs occur when models produce correct-sounding but factually incorrect outputs, posing challenges for ensuring trustworthy AI." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

### NEW QUESTION # 17

Which technique is designed to train a deep learning model by adjusting the weights of the neural network based on the error between the predicted and actual outputs?

- A. Gradient Boosting
- B. Backpropagation
- C. K-means Clustering
- D. Principal Component Analysis

**Answer: B**

Explanation:

Backpropagation is a fundamental technique in training deep learning models, as emphasized in NVIDIA's Generative AI and LLMs course. It is designed to adjust the weights of a neural network by propagating the error between the predicted and actual outputs backward through the network. This process calculates gradients of the loss function with respect to each weight using the chain rule, enabling iterative weight updates via gradient descent to minimize the error. Backpropagation is essential for optimizing neural networks, including those used in large language models (LLMs), by fine-tuning weights to improve predictions. Option A, Gradient Boosting, is incorrect as it is an ensemble method for decision trees, not neural networks. Option B, Principal Component Analysis, is a dimensionality reduction technique, not a training method. Option C, K-means Clustering, is an unsupervised clustering algorithm, unrelated to supervised weight adjustment. The course highlights: "Backpropagation is used to train neural networks by computing gradients of the loss function and updating weights to minimize prediction errors, a critical process in deep learning models like Transformers." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

### NEW QUESTION # 18

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

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

**Answer: A**

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 # 19

In the context of transformer-based large language models, how does the use of layer normalization mitigate the challenges associated with training deep neural networks?

- A. It replaces the attention mechanism to improve sequence processing efficiency.
- B. It reduces the computational complexity by normalizing the input embeddings.
- C. It increases the model's capacity by adding additional parameters to each layer.
- D. It stabilizes training by normalizing the inputs to each layer, reducing internal covariate shift.

**Answer: D**

Explanation:

Layer normalization is a technique used in transformer-based large language models (LLMs) to stabilize and accelerate training by normalizing the inputs to each layer. According to the original transformer paper ("Attention is All You Need," Vaswani et al., 2017) and NVIDIA's NeMo documentation, layer normalization reduces internal covariate shift by ensuring that the mean and variance of activations remain consistent across layers, mitigating issues like vanishing or exploding gradients in deep networks. This is particularly crucial in transformers, which have many layers and process long sequences, making them prone to training instability. By normalizing the activations (typically after the attention and feed-forward sub-layers), layer normalization improves gradient flow and convergence. Option A is incorrect, as layer normalization does not reduce computational complexity but adds a small overhead. Option C is false, as it does not add significant parameters. Option D is wrong, as layer normalization complements, not replaces, the attention mechanism.

References:

Vaswani, A., et al. (2017). "Attention is All You Need."

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

### NEW QUESTION # 20

.....

Are you aware of the importance of the NCA-GENL certification? If your answer is not, you may place yourself at the risk of being eliminated by the labor market. Because more and more companies start to pay high attention to the ability of their workers, and the NCA-GENL Certification is the main reflection of your ability. And our NCA-GENL exam question are the right tool to help you get the certification with the least time and efforts. Just have a try, then you will love them!

**Test NCA-GENL Assessment:** <https://www.exams4collection.com/NCA-GENL-latest-braindumps.html>

NVIDIA NCA-GENL Trustworthy Exam Torrent You can be confident to them who know exactly what is going to be in your real test, we believe you can make it with the help of it, However, not only the good start can Test NCA-GENL Assessment - NVIDIA

These movies have stories to them. The business representative, NCA-GENL New Exam Brainsdumps Betty, presented the two of them with a business rule for giving discounts that she had obtained from one of the stakeholders.

You can be confident to them who know exactly what is going to be in your real NCA-GENL Trustworthy Exam Torrent test, we believe you can make it with the help of it, However, not only the good start can NVIDIA Generative AI LLMs free download pdf provide you but also the good ending.

[illegible]

2026 Latest Exams4Collection NCA-GENL PDF Dumps and NCA-GENL Exam Engine Free Share:  
<https://drive.google.com/open?id=1b0flGCbm4jsEWFbQdbOwa8tOsHJZTlpH>