

NVIDIA NCA-GENL受験対策解説集: NVIDIA Generative AI LLMs - Topexam正確な関連合格問題無料ダウンロード



P.S. TopexamがGoogle Driveで共有している無料かつ新しいNCA-GENLダンプ: https://drive.google.com/open?id=1t60LJmfXnRtJR2M54qwZtxs2WsSPUR_F

NCA-GENL問題集を買うとき、支払いが成功したら、お客様は問題集をダウンロードできます。NCA-GENL問題集の有効性を確保する為に、NVIDIAはNCA-GENL問題集のに対して、定期的に検査します。そうすれば、お客様にNCA-GENL問題集の最新版を提供できます。

NVIDIA認定を取得したい場合は、行動し始めてみませんか？最初のステップは、NCA-GENL試験に合格することです。時間は誰も待っていません。NCA-GENL試験に合格した場合にのみ、より良いプロモーションを取得できます。そして、あなたがより効率的にそれを渡したいなら、私たちはあなたにとって最高のパートナーでなければなりません。私たちはプロのNCA-GENL質問トレントプロバイダーであり、NCA-GENLトレーニング資料は信頼に値します。NCA-GENLラーニングガイドに多大な努力を払っているため、10年以上にわたってこの分野でより良い成果を上げています。NCA-GENL学習ガイドが最適です。

>> NCA-GENL受験対策解説集 <<

効果的なNCA-GENL受験対策解説集試験-試験の準備方法-高品質なNCA-GENL関連合格問題

NVIDIA NCA-GENL資格認定はIT技術領域に従事する人に必要があります。我々社のNVIDIA NCA-GENL試験練習問題はあなたに試験うま合格できるのを支援します。あなたの取得したNVIDIA NCA-GENL資格認定は、工作中に核心技術知識を同僚に認可されるし、あなたの技術信頼度を増強できます。

NVIDIA Generative AI LLMs 認定 NCA-GENL 試験問題 (Q30-Q35):

質問 # 30

What is 'chunking' in Retrieval-Augmented Generation (RAG)?

- A. A method used in RAG to generate random text.
- B. Rewrite blocks of text to fill a context window.
- C. A concept in RAG that refers to the training of large language models.
- **D. A technique used in RAG to split text into meaningful segments.**

正解: D

解説:

Chunking in Retrieval-Augmented Generation (RAG) refers to the process of splitting large text documents into smaller, meaningful segments (or chunks) to facilitate efficient retrieval and processing by the LLM.

According to NVIDIA's documentation on RAG workflows (e.g., in NeMo and Triton), chunking ensures that retrieved text fits within the model's context window and is relevant to the query, improving the quality of generated responses. For example, a long document might be divided into paragraphs or sentences to allow the retrieval component to select only the most pertinent chunks. Option A is incorrect because chunking does not involve rewriting text. Option B is wrong, as chunking is not about generating random text. Option C is unrelated, as chunking is not a training process.

References:

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

Lewis, P., et al. (2020). "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks."

質問 # 31

Which metric is commonly used to evaluate machine-translation models?

- A. F1 Score
- **B. ROUGE score**
- C. Perplexity
- D. BLEU score

正解: B

解説:

The BLEU (Bilingual Evaluation Understudy) score is the most commonly used metric for evaluating machine-translation models. It measures the precision of n-gram overlaps between the generated translation and reference translations, providing a quantitative measure of translation quality. NVIDIA's NeMo documentation on NLP tasks, particularly machine translation, highlights BLEU as the standard metric for assessing translation performance due to its focus on precision and fluency. Option A (F1 Score) is used for classification tasks, not translation. Option C (ROUGE) is primarily for summarization, focusing on recall.

Option D (Perplexity) measures language model quality but is less specific to translation evaluation.

References:

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

Papineni, K., et al. (2002). "BLEU: A Method for Automatic Evaluation of Machine Translation."

質問 # 32

What is confidential computing?

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

正解: C

解説:

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.

質問 # 33

You have access to training data but no access to test data. What evaluation method can you use to assess the performance of your AI model?

- A. Average entropy approximation
- B. Greedy decoding
- **C. Cross-validation**
- D. Randomized controlled trial

正解: C

解説:

When test data is unavailable, cross-validation is the most effective method to assess an AI model's performance using only the training dataset. Cross-validation involves splitting the training data into multiple subsets (folds), training the model on some folds, and validating it on others, repeating this process to estimate generalization performance. NVIDIA's documentation on machine learning workflows, particularly in the NeMo framework for model evaluation, highlights k-fold cross-validation as a standard technique for robust performance assessment when a separate test set is not available. Option B (randomized controlled trial) is a clinical or experimental method, not typically used for model evaluation. Option C (average entropy approximation) is not a standard evaluation method. Option D (greedy decoding) is a generation strategy for LLMs, not an evaluation technique.

References:

NVIDIA NeMo Documentation: https://docs.nvidia.com/deeplearning/nemo/user-guide/docs/en/stable/nlp/model_finetuning.html
Goodfellow, I., et al. (2016). "Deep Learning." MIT Press.

質問 # 34

Your company has upgraded from a legacy LLM model to a new model that allows for larger sequences and higher token limits. What is the most likely result of upgrading to the new model?

- A. The number of tokens is fixed for all existing language models, so there is no benefit to upgrading to higher token limits.
- **B. The newer model allows larger context, so outputs will improve, but you will likely incur longer inference times.**
- C. The newer model allows the same context lengths, but the larger token limit will result in more comprehensive and longer outputs with more detail.
- D. The newer model allows for larger context, so the outputs will improve without increasing inference time overhead.

正解: B

解説:

Upgrading to a new LLM with larger sequence lengths and higher token limits, as discussed in NVIDIA's Generative AI and LLMs course, typically allows the model to process larger contexts, leading to improved output quality due to better understanding of extended dependencies in text. However, handling larger sequences increases computational requirements, often resulting in longer inference times, especially on the same hardware. This trade-off is a key consideration in LLM deployment. Option A is incorrect, as token limits vary across models, and higher limits offer benefits. Option B is wrong, as larger context processing typically increases inference time. Option C is inaccurate, as higher token limits primarily enable larger context, not just longer outputs. The course notes: "Larger sequence lengths in LLMs allow for improved output quality by capturing more context, but this often comes at the cost of increased inference times due to higher computational demands." References: NVIDIA Building Transformer-Based Natural Language Processing Applications course; NVIDIA Introduction to Transformer-Based Natural Language Processing.

質問 # 35

.....

TopexamのNCA-GENL問題集はあなたを楽に試験の準備をやらせます。それに、もし最初で試験を受ける場合、試験のソフトウェアのバージョンを使用することができます。これは完全に実際の試験雰囲気とフォーマットをシミュレートするソフトウェアですから。このソフトで、あなたは事前に実際の試験を感じることができます。そうすれば、実際のNCA-GENL試験を受けるときに緊張をすることは不会です。ですから、心のリラック

NCA-GENL関連合格問題: https://www.topexam.jp/NCA-GENL_shiken.html

むしろそうなるのが怖くて動画は匿名で公開していただくのだ、新しい生NCA-GENL命も授かった、試験の準備は時間とエネルギーがかかります、これは受験生の皆様を助けた結果です、自分で試してみれば、弊社は信用できると分かります。

[illegible]

2026年Topexamの最新NCA-GENL PDFダンプおよびNCA-GENL試験エンジンの無料共有: https://drive.google.com/open?id=1t60IJmfXnRtJR2M54qwZtxs2WsSPUR_F