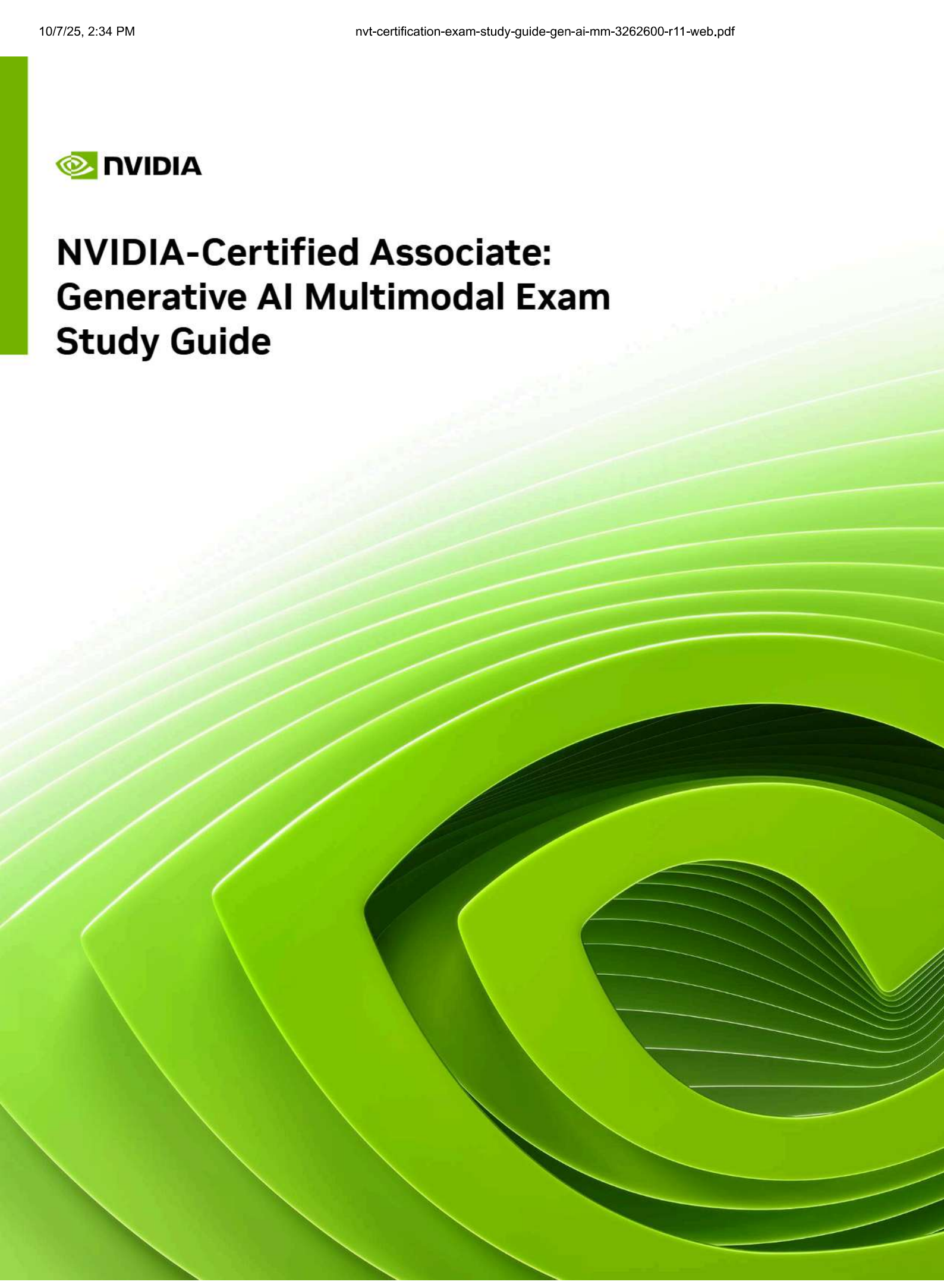
# NVIDIA-Certified Associate: Generative AI Multimodal Exam Study Guide



# NVIDIA-Certified Associate: Generative AI Multimodal Exam Study Guide

Contents

This study guide provides an overview of each topic covered on the NVIDIA Generative AI Multimodal certification exam, recommended training, and suggested reading to prepare for the exam.

Information about NVIDIA certifications can be found here.

## Job Description

The generative AI multimodal associate developer is responsible for contributing to the development, programming, and quality assurance of state-of-the-art generative AI multimodal models. They work with a team of skilled AI professionals to develop datasets, select models to train, train models, and implement model testing and debugging processes. They should have an understanding of the deployment of models for applications. They'll also be responsible for developing high-quality software design and construction, programming in a variety of languages and platforms, and maintaining system updates.

## Job Responsibilities

1. Collaborate with the AI development team to design, code, test, debug, and document programming applications.
2. Perform system analysis to ensure software and systems meet required specifications.
3. Aid in integrating new AI language models into existing systems or creating new ones as needed.
4. Assist in the assessment and resolution of application and system performance issues.
5. Stay updated on new AI models and other developments related to language learning models.
6. Contribute to the production of technical documents and manuals.
7. Conduct software programming and documentation development under the direction of senior staff.
8. Perform prompt engineering.
9. Select models.
10. Define, curate, and annotate/label multimedia (audio, video, images) datasets.
11. Perform experimentations like A/B testing, evaluating prompts, evaluating models, and producing POCs.

## Recommended Qualifications and Experience

1. Bachelor's degree in computer science, software engineering, AI, or a related field
2. Knowledge of Python, C, and AI frameworks (PyTorch, TensorFlow, etc.)
3. Solid understanding of neural networks and deep learning models

# Certification Topics and References

## Core Machine Learning and AI Knowledge: Exam Weight 20%

| | |
|---|---|
| | Knowledge of algorithms, conventions, and techniques that allow computers to learn from and make predictions or decisions based on data. |
| 1.1 | Control stability of training in multimodal settings |
| 1.2 | Develop content for introduction to multimodal loss functions. |
| 1.3 | Familiarity with fundamentals of machine learning (e.g., feature engineering, model comparison, cross validation). |
| 1.4 | Understand nonsequential neural networks and residual connections. |
| 1.5 | Design statistical analysis for evaluating multimodal pipelines. |
| 1.6 | Develop content for multimodal-specific transfer learning. |
| 1.7 | Familiarity with emerging multimodal trends and technologies. |
| 1.8 | Contribute to the design, development, and deployment of energy-efficient and trustworthy multimodal AI models. |
| 1.9 | Use prompt engineering principles to create prompts to achieve desired results. |
| 1.10 | Understand deep learning frameworks such as TensorFlow or PyTorch. |

### Recommended Training (Optional)

Course reference: Getting Started With Deep Learning, Fundamentals of Deep Learning
- Learn the fundamental techniques and tools required to train a deep learning model.
- Gain experience with common deep learning data types and model architectures.

Course reference: Introduction to Transformer-Based Natural Language Processing
- Learn how transformers are used as the building blocks of modern large language models (LLMs).

Course reference: Building Transformer-Based Natural Language Processing Applications
- Learn how transformers are used as the basic building blocks of modern LLMs for natural language processing (NLP) applications.
- Learn how self-supervision improves upon the transformer architecture in BERT, Megatron, and other LLM variants for superior NLP results.

Course reference: Building AI Agents With Multimodal Models
- Learn about different data types and how to make them neural-network ready.
- Learn about model fusion and the differences between early, late, and intermediate fusion.
- Learn the difference between modality and agent orchestration.

### Suggested Readings

- A Complete Machine Learning Project Walk-Through in Python: Part One, by Will Koehrsen, Towards Data Science
- Overfit and Underfit, TensorFlow Core
- Image Classification, NVIDIA NGC™
- What Is a GAN?—Generative Adversarial Networks Explained
- Intuitively Understanding Convolutions for Deep Learning, by Irhum Shafkat, Towards Data Science
- Multimodal Machine Learning
- What Is Overfitting in Deep Learning [+10 Ways to Avoid It]
- Introduction to Diffusion Models for Machine Learning

## **Data Analysis:** Exam Weight 10%

| |
|---|
| Inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. |
| 2.1 Awareness of the process of extracting insights from large datasets using data mining, data visualization, and similar techniques. |
| 2.2 Develop content for attention maps in multimodal settings. |
| 2.3 Create graphs, charts, or other visualizations to convey the results of data analysis using specialized software. |
| 2.4 Identify relationships and trends or any factors that could affect the results of research. |

### Recommended Training (Optional)

Course reference: Getting Started With Deep Learning, Fundamentals of Deep Learning

- Enhance datasets through data augmentation to improve model accuracy.

Course reference: Introduction to Transformer-Based Natural Language Processing

- Use transformer-based models for various NLP tasks, including text classification, named-entity recognition (NER), author attribution, and question answering.

Course reference: Building Transformer-Based Natural Language Processing Applications

- Leverage pretrained, modern LLM models to solve multiple NLP tasks such as text classification, named-entity recognition (NER), and question answering.

Course reference: Building AI Agents With Multimodal Models

- Learn about PDF extraction using OCR.

### Suggested Readings

- What Is Data Visualization? Definition, Examples, and Learning Resources
- Essential Chart Types for Data Visualization, Atlassian
- 7 Ways to Handle Missing Values in Machine Learning, by Satyam Kumar, Towards Data Science
- Guide to Data Cleaning: Definition, Benefits, Components, and How to Clean Your Data

## **Experimentation:** Exam Weight 25%

The study of how to perform, evaluate, and interpret experiments, including AI model evaluation and the evaluation of various model architectures.

| | |
|---|---|
| 3.1 | Assist in developing and testing multimodal AI models. |
| 3.2 | Manage and preprocess data from various sources. |
| 3.3 | Use multimodal models to improve explainability. |
| 3.4 | Test data quality and consistency in a multimodal setting. |
| 3.5 | Test AI models to ensure their accuracy and effectiveness. |

### Recommended Training (Optional)

Course reference: Getting Started With Deep Learning, Fundamentals of Deep Learning
- Enhance datasets through data augmentation to improve model accuracy.

Course reference: Building Transformer-Based Natural Language Processing Applications
- Leverage pretrained, modern LLM models to solve multiple NLP tasks such as text classification, named-entity recognition (NER), and question answering.

Course reference: Building Conversational AI Applications
- Customize and deploy automatic speech recognition (ASR) and text-to-speech (TTS) models on NVIDIA® Riva.
- Build and deploy an end-to-end conversational AI pipeline, including ASR, NLP, and TTS models on Riva.
- Deploy a production-level conversational AI application with a Helm chart for scaling in Kubernetes clusters.

Course reference: Generative AI With Diffusion Models
- Improve the quality of generated images with the denoising diffusion process.
- Control the image output with context embeddings.

### Suggested Readings

- A Complete Guide to Data Augmentation, DataCamp
- Introduction to Large Language Models and the Transformer Architecture
- CLIP: Connecting Text and Images
- Basics of Speech Recognition and Customization of Riva ASR
- Law Professor Explores Racial Bias Implications in Facial Recognition Technology, University of Calgary
- GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium

## **Multimodal Data:** Exam Weight 15%

Multimodal data involves the integration, curation, and quality assessment of diverse data types such as text, images, audio, time-series, and geospatial information, while also addressing challenges related to missing or incomplete information across these different modalities.

| | |
|---|---|
| 4.1 | Assist in the deployment and evaluations of model scalability, performance, and reliability under the supervision of senior team member. |
| 4.2 | Build LLM use cases such as retrieval-augmented generation (RAG), chatbots, and summarizers. |
| 4.3 | Familiarity with the capabilities of Python natural language packages (spaCy, NumPy, vector databases, etc.). |
| 4.4 | Identify system data, hardware, or software components required to meet user needs. |
| 4.5 | Monitor the functioning of data collection, experiments, and other software processes. |
| 4.6 | Use Python packages (spaCy, NumPy, Keras, etc.) to implement specific traditional machine learning analyses. |
| 4.7 | Write software components or scripts under the supervision of a senior team member. |

### Recommended Training (Optional)

Course reference: Building Conversational AI Applications
- Customize and deploy ASR and TTS models on Riva.
- Build and deploy an end-to-end conversational AI pipeline, including ASR, NLP, and TTS models on Riva.

Course reference: Generative AI With Diffusion Models
- Generate images from English text prompts using Contrastive Language-Image Pretraining (CLIP).

Course reference: Building AI Agents With Multimodal Models
- Learn about model fusion and the differences between early, late, and intermediate fusion.
- Learn the difference between modality and agent orchestration.

### Suggested Readings

- Multimodal Machine Learning
- CLIP: Connecting Text and Images
- Multimodal Machine Learning: Data Fusion, by Adrienne Kline, Towards AI
- Anomaly Detection With Auto-Encoders
- Effective Techniques for Multimodal Data Fusion: A Comparative Analysis, PubMed Central (PMC)

## **Performance Optimization:** Exam Weight 10%

Performance optimization in AI entails refining multimodal AI models for energy efficiency, trustworthiness, and accuracy through design contributions, transfer learning content development, supervised training enhancements, hyperparameter tuning, rigorous testing, and computational advancements.

| | |
|---|---|
| 5.1 | Enhance computational efficiency and improve the accuracy of outputs in AI models. |
| 5.2 | Optimize the performance of AI models, including tuning hyperparameters. |
| 5.3 | Develop content for multimodal-specific transfer learning. |
| 5.4 | Assist in model training and training optimization under the supervision of a senior team member. |

### Recommended Training (Optional)

Course reference: Getting Started With Deep Learning, Fundamentals of Deep Learning
> Leverage transfer learning between models to achieve efficient results with less data and computation.

Course reference: Building Transformer-Based Natural Language Processing Applications
> Manage inference challenges and deploy refined models for live applications.

Course reference: Building Conversational AI Applications
> Deploy a production-level conversational AI application with a Helm chart for scaling in Kubernetes clusters.

### Suggested Readings

> Mixed Precision Training
> An Intuitive Explanation of Connectionist Temporal Classification, by Harald Scheidl, Towards Data Science
> Basics of Quantization in Machine Learning (ML) for Beginners
> What Is Energy Efficiency and Why Is It Important?, NVIDIA
> Neural Network Pruning With Combinatorial Optimization

## **Software Development:** Exam Weight 15%

Design and implement neural network architectures, such as U-Nets for generative image tasks, integrate text-to-image AI models like CLIP, and apply prompt engineering to refine and direct the generative capabilities of these systems. Includes familiarity with NVIDIA SDKs such as Riva, NeMo™, Triton™, and Avatar Cloud Engine (ACE).

| | |
|---|---|
| 6.1 | Collaborate with the client during requirements acquisition, data gathering, progress reporting, deployment, and integration. |
| 6.2 | Ensure adherence to best practices and maintain high standards of software quality and reliability. |
| 6.3 | Use prompt engineering to better influence the output of generative AI models. |
| 6.4 | Build a U-Net to generate images from pure noise and as a type of autoencoder. |
| 6.5 | Generate images from English text prompts using CLIP, and use CLIP to train a text-to-image diffusion model. |

### Recommended Training (Optional)

Course reference: Getting Started With Deep Learning, Fundamentals of Deep Learning
> Build confidence to take on your own project with a modern deep learning framework.

Course reference: Generative AI With Diffusion Models
> Generate images from pure noise.
> Generate images from English text prompts using CLIP.

Course reference: Building AI Agents With Multimodal Models
> Customize NVIDIA AI Blueprints with VIA.

### Suggested Readings

> Zero-Shot Image Classification With OpenAI's CLIP, Pinecone
> CUDA Deep Neural Network (cuDNN), NVIDIA Developer
> Transfer Learning, Wikipedia
> TensorFlow, NVIDIA NGC
> What Is the Importance of A/B Testing in Machine Learning?
> CLIP: Connecting Text and Images

## **Trustworthy AI:** Exam Weight 5%

Create and assess ethical, energy-conscious, and reliable artificial intelligence systems that are capable of interpreting and integrating various forms of data, ensuring that they're designed and applied in a manner that's transparent, fair, and verifiable.

| | |
|---|---|
| 7.1 | Describe the ethical principles of trustworthy AI. |
| 7.2 | Describe the balance between data privacy and the importance of data consent. |
| 7.3 | Describe how to use NVIDIA and other technologies to improve AI trustworthiness. |
| 7.4 | Describe how to minimize bias in AI systems. |

### Recommended Training (Optional)

Generative AI With Diffusion Models

> Learn about content authenticity and how to build trustworthy models.

### Suggested Readings

> Trustworthy AI for a Better World, NVIDIA
> What Is Trustworthy AI?, NVIDIA Blog
> What Is Retrieval-Augmented Generation aka RAG, NVIDIA Blog

## Questions?

Contact us here.