Generative AI Innovative Applications

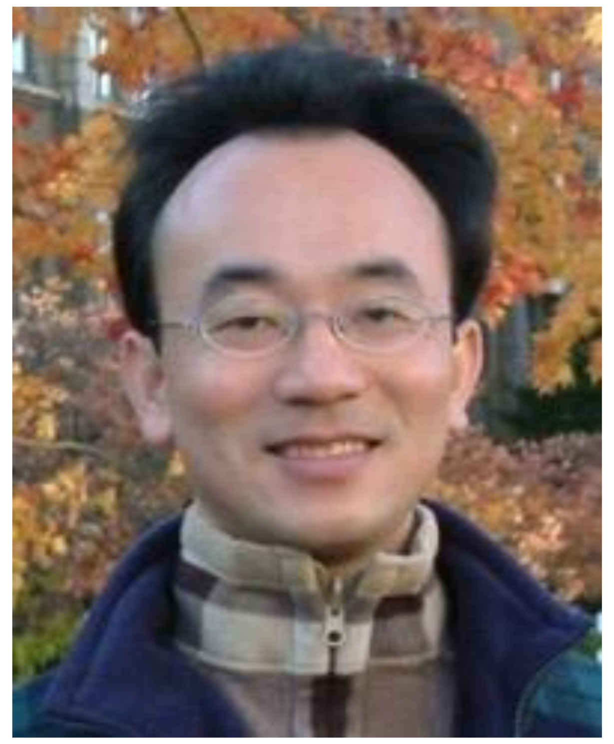
國立臺北大學
National Taipei University

NVIDIA Generative AI with Diffusion Models
Part I

1132GAIIA07
MBA, IM, NTPU (M6031) (Spring 2025)
Tue 2, 3, 4 (9:10-12:00) (B3F17)

https://meet.google.com/paj-zhhj-mya

NVIDIA
aws educate | Cloud Ambassador
2020 Cohort
aws academy Accredited Educator
aws certified Cloud Practitioner
aws certified Solutions Architect Associate



Min-Yuh Day, Ph.D, Professor

Institute of Information Management, National Taipei University

https://web.ntpu.edu.tw/~myday

2025-04-22

Syllabus

Week Date Subject/Topics

1 2025/02/18 Introduction to Generative AI Innovative Applications

2 2025/02/25 Transformers for Natural Language Processing and Computer Vision

3 2025/03/04 Large Language Models (LLMs), NVIDIA Building RAG Agents with LLMs Part I

4 2025/03/11 Case Study on Generative AI Innovative Applications I

5 2025/03/18 NVIDIA Building RAG Agents with LLMs Part II

6 2025/03/25 NVIDIA Building RAG Agents with LLMs Part III

Syllabus

Week	Date	Subject/Topics
7	2025/04/01	Self-Learning
8	2025/04/08	Midterm Project Report
9	2025/04/15	Generative AI for Multimodal Information Generation
10	2025/04/22	NVIDIA Generative AI with Diffusion Models Part I
11	2025/04/29	NVIDIA Generative AI with Diffusion Models Part II
12	2025/05/06	Case Study on Generative AI Innovative Applications II

Syllabus

Week Date Subject/Topics

13 2025/05/13 NVIDIA DLI and NVIDIA-Certified Associate (NCA)
Generative AI LLMs Certification [Invited Talk]

14 2025/05/20 NVIDIA Generative AI with Diffusion Models Part III
AI Agents and Large Multimodal Agents (LMAs)

15 2025/05/27 Final Project Report I

16 2025/06/03 Final Project Report II

NVIDIA
Generative AI with Diffusion Models
Part I
[0,1,2] [3, 4] [5, 6]

Generative AI with Diffusion Models

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- Part 3: Optimizations
- Part 4: Classifier-Free Diffusion Guidance (CFG)
- Part 5: CLIP (Contrastive Language-Image Pre-Training)
- Part 6: Wrap-up & Assessment

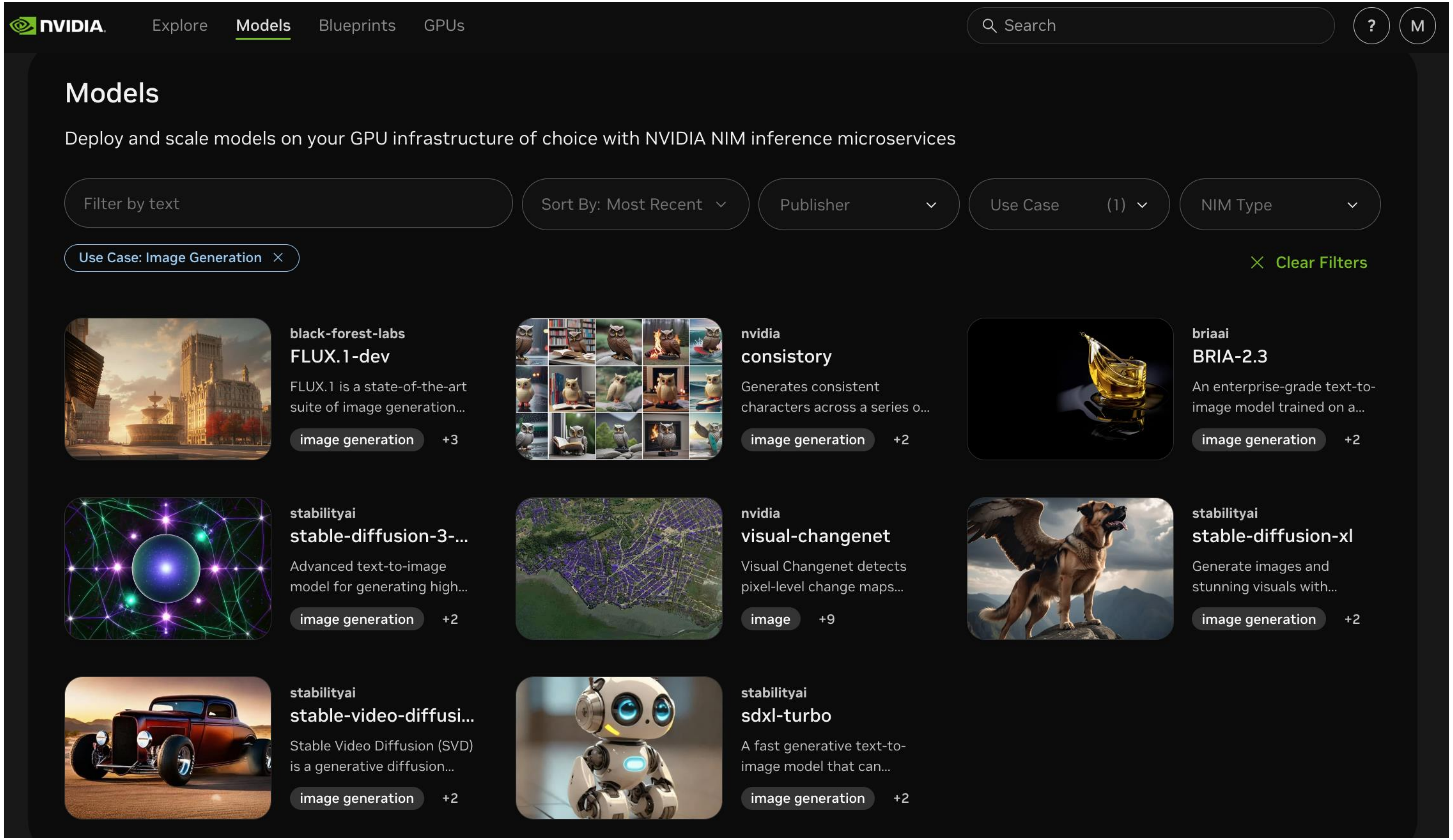
https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Diffusion Models

diffusion models are trained to denoise noisy images,
and can generate images by iteratively denoising pure noise.

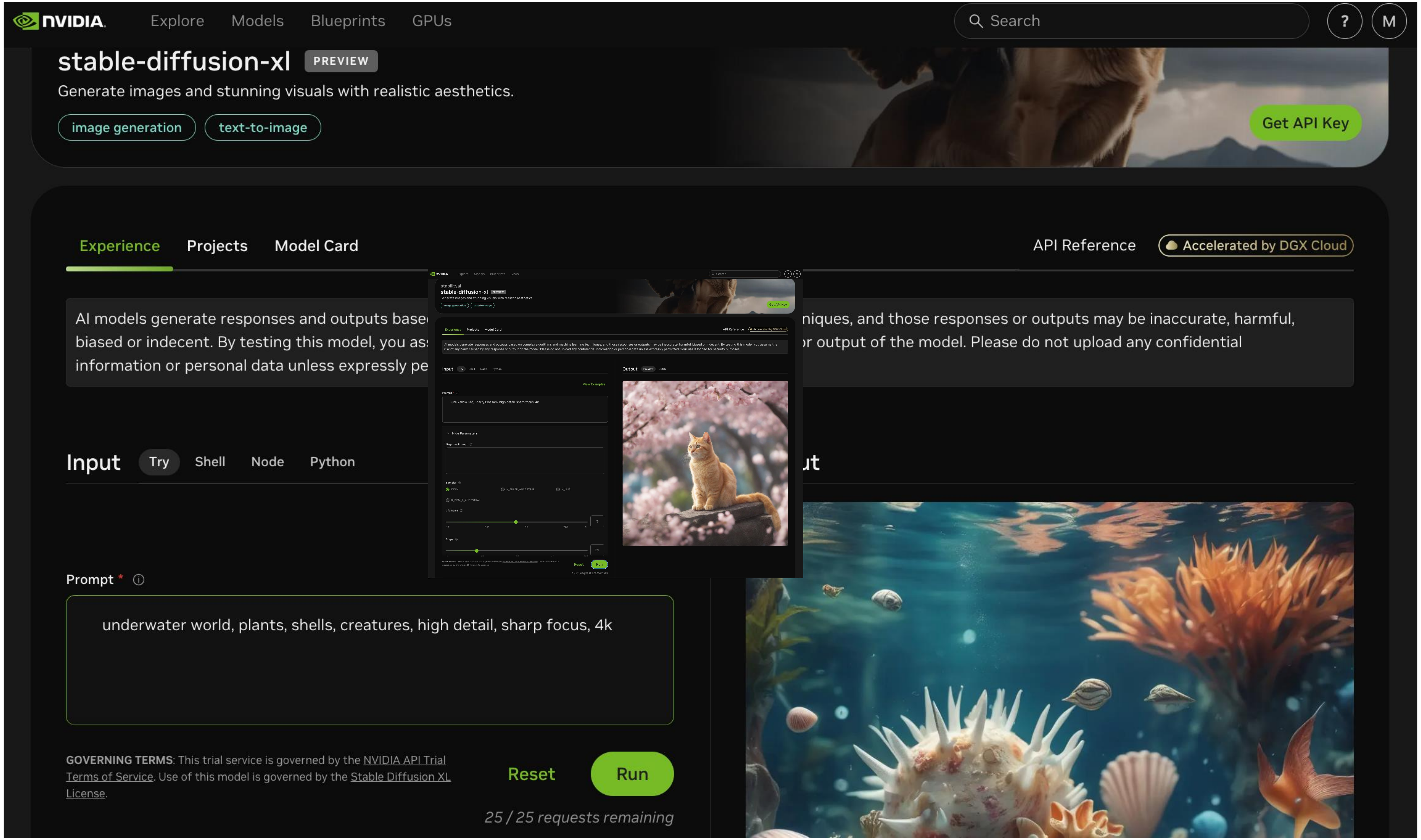
Source: https://keras.io/examples/generative/ddim/

build.nvidia.com **Models: Image Generation**



https://build.nvidia.com/models?filters=usecase%3Ausecase_image_gen

build.nvidia.com **stabilityai/stable-diffusion-xl**

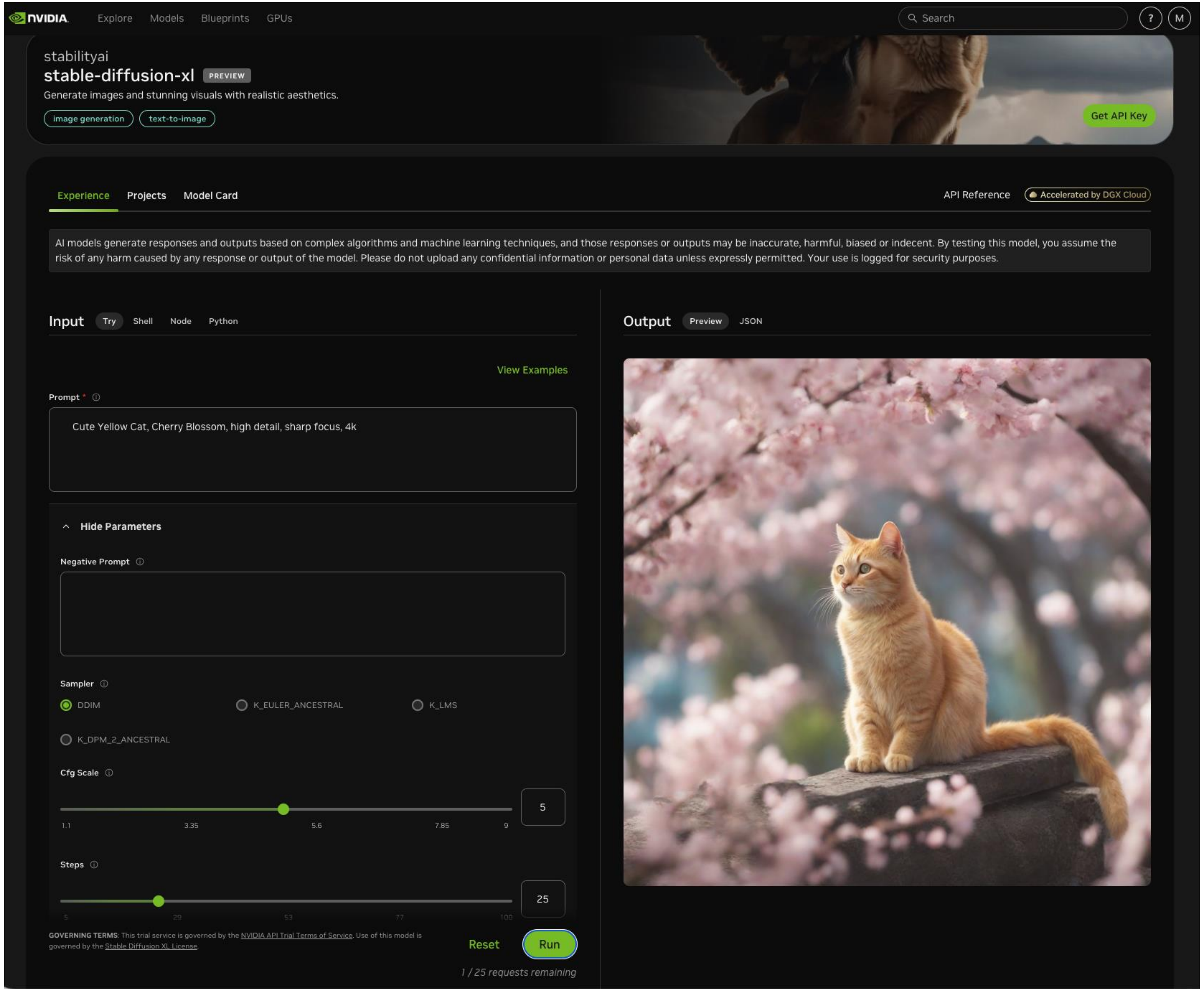


https://build.nvidia.com/stabilityai/stable-diffusion-xl

build.nvidia.com **stabilityai/stable-diffusion-xl**

Input Prompt

Cute Yellow Cat, Cherry Blossom, high detail, sharp focus, 4k



https://build.nvidia.com/stabilityai/stable-diffusion-xl

Spring 2025

Generative AI Innovative Applications

NVIDIA

University Ambassador

This certificate acknowledges that

Min-Yuh Day

has been certified to deliver NVIDIA instructor-led workshop for academia

Greg Estes

Vice President, NVIDIA

Issue Date: : March 7, 2025

Ambassador Certification ID: cCFh1ZWWTvqKTq7dcKkEWw

https://learn.nvidia.com/certificates?id=cCFh1ZWWTvqKTq7dcKkEWw

NVIDIA

Certified Instructor

This certificate acknowledges that

Min-Yuh Day

has been certified to deliver the instructor-led workshop

Building RAG Agents with LLMs

Greg Estes

Vice President, NVIDIA

Issue Date: : March 7, 2025

Certification ID: OVmqY4cSSya0BdMQBWHxzw

https://learn.nvidia.com/certificates?id=OVmqY4cSSya0BdMQBWHxzw

NVIDIA

Certificate of Completion

This certificate is awarded to

Min-Yuh Day

for successfully completing

Building RAG Agents with LLMs

Greg Estes

Vice President, NVIDIA

Issue Date: : December 8, 2024

Certification ID: ed-qOCIMQaatzU8SNUNxgw |

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw/courses/course?course_id=course-v1:DLI+S-FX-15+V1

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw

NVIDIA

Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of

Generative AI with Diffusion Models

Greg Estes

Vice President, NVIDIA

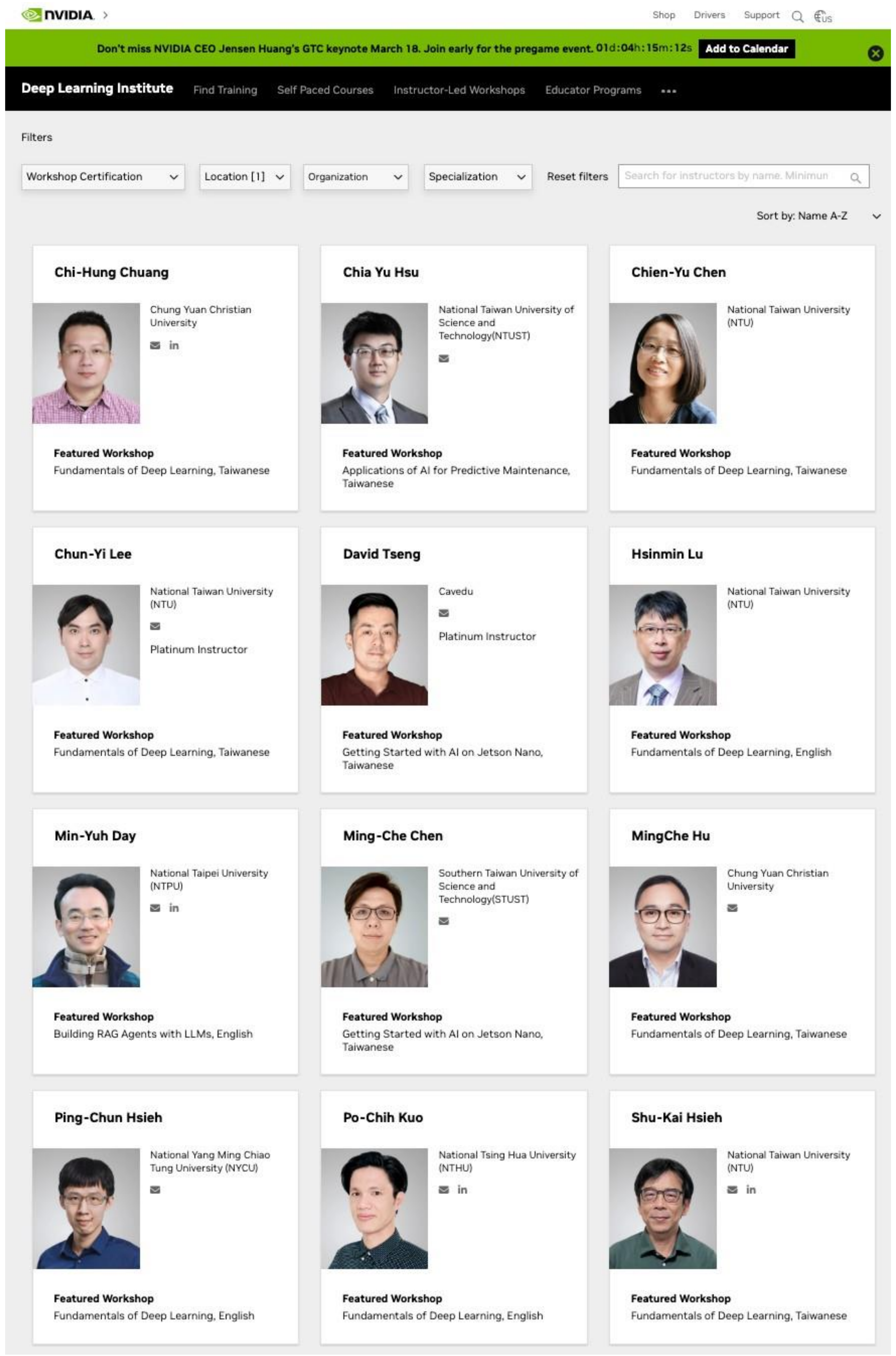
Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

https://learn.nvidia.com/certificates?id=q3oo-oBhTQKtyCCote2E-Q

NVIDIA Certified Instructors

(12 from Taiwan)

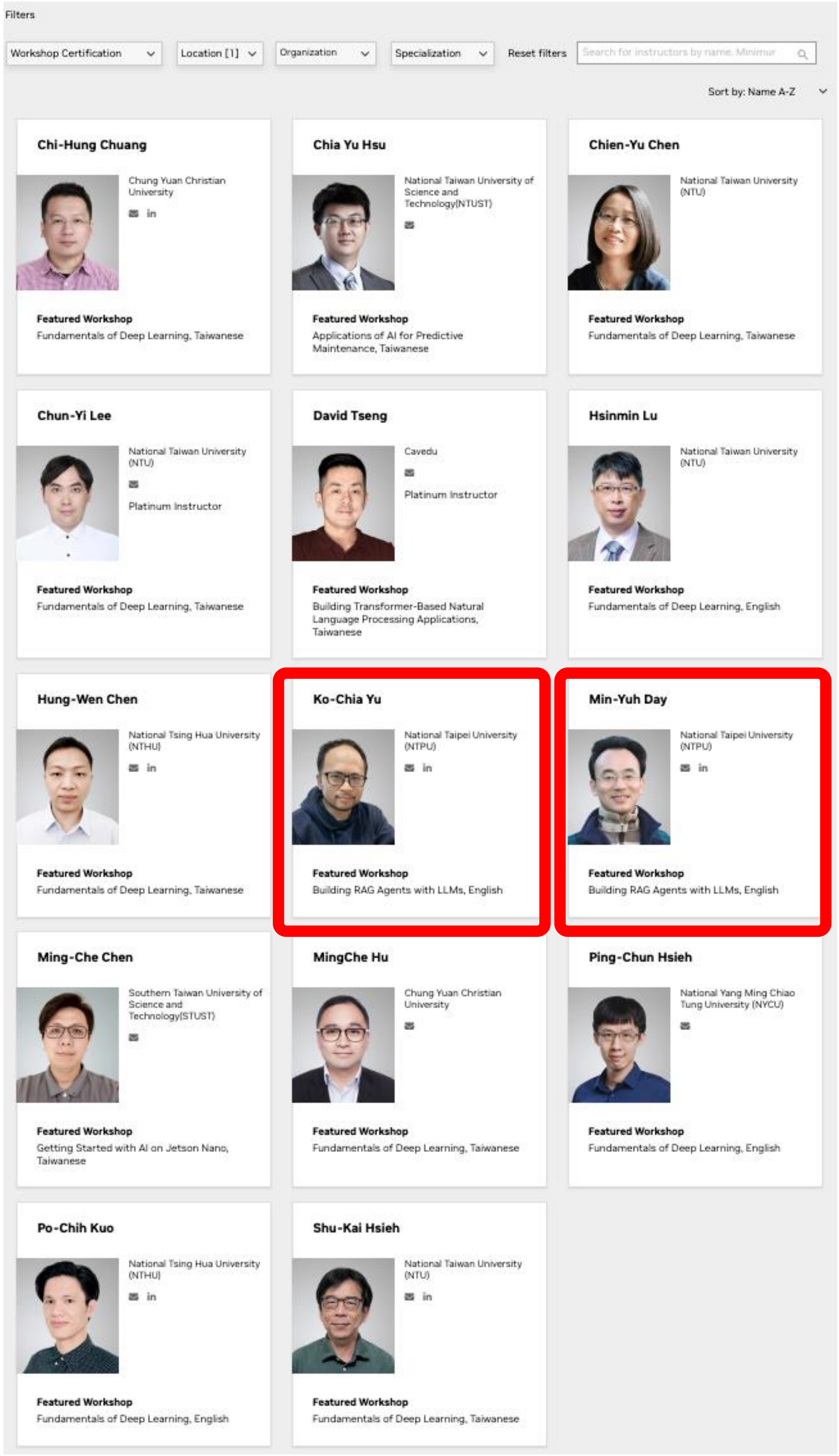


https://www.nvidia.com/en-us/training/instructor-directory/search/

NVIDIA Certified Instructors

(14 from Taiwan)

(2 from NTPU)(April 2025)



https://www.nvidia.com/en-us/training/instructor-directory/search/

NVIDIA Developer Program

https://developer.nvidia.com/join-nvidia-developer-program

NVIDIA
Deep Learning Institute (DLI)

https://learn.nvidia.com/

Get NVIDIA DLI Certificate

- **Step 1. Join NVIDIA Developer Program (Free)**
  https://developer.nvidia.com/join-nvidia-developer-program
- **Step 2. Visit NVIDIA Deep Learning Institute (DLI)**
  https://learn.nvidia.com/
- **Step 3. Enroll "Generative AI with Diffusion Models"**
  Self-Paced Course ($90)
  https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA

Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to $90)

Generative AI and LLMs | Graphics and Simulation | Accelerated Computing | Data Science | Deep Learning

8 hours

Getting Started With Deep Learning

Explore the fundamentals of deep learning by training neural networks and using results to improve performance and capabilities.

2 hours

Modeling Time-Series Data With Recurrent Neural Networks in Keras

Explore how to classify and forecast time-series data using recurrent neural networks (RNNs), such as modeling a patient's health over time.

4 hours

Deploying a Model for Inference at Production Scale

Learn how to deploy your own machine learning models on a GPU server.

8 hours

Building Real-Time Video AI Applications

Gain the knowledge and skills needed to enable the real-time transformation of raw video data from widely deployed camera sensors into deep learning-based insights.

2 hours

Introduction to Graph Neural Networks

Learn the basic concepts, models, and applications of graph neural networks.

4 hours

Introduction to Physics-Informed Machine Learning With Modulus

Learn the various building blocks of NVIDIA Modulus, which turbocharges use cases by building physics-based deep learning models that are 100,000X faster than traditional methods and offers high-fidelity simulation results.

2 hours

Get Started With Highly Accurate Custom ASR for Speech AI

Learn to build, train, fine-tune, and deploy a GPU-accelerated automatic speech recognition (ASR) service with NVIDIA® Riva that includes customized features.

2 hours

Integrating Sensors With NVIDIA DRIVE

Find out how to integrate automotive sensors into your applications using NVIDIA DRIVE®.

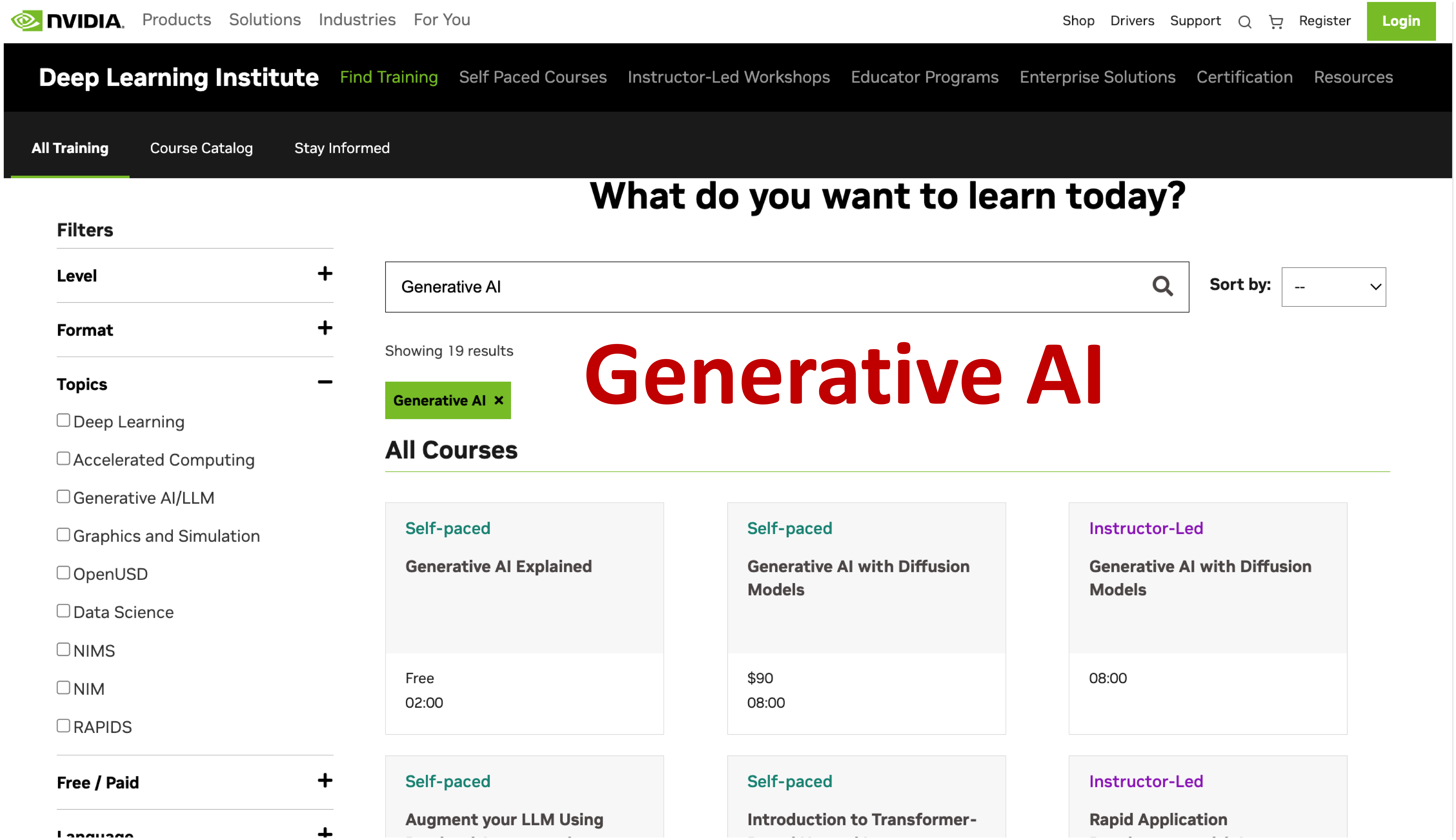
https://developer.nvidia.com/join-nvidia-developer-program

NVIDIA Deep Learning Institute (DLI)

Self-Paced Course **Generative AI Explained** Free 2 hours	Self-Paced Course **Getting Started With Deep Learning** Certificate available $90 8 hours	Instructor-Led Workshop **Fundamentals of Deep Learning** Certificate available $500 8 hours	Self-Paced Course **Introduction to Transformer-Based Natural Language Processing** Certificate available $30 6 hours
Self-Paced Course **Building RAG Agents With LLMs** Certificate available Free 8 hours	Instructor-Led Workshop **Building RAG Agents With LLMs** Certificate available $500 8 hours	Self-Paced Course **Generative AI with Diffusion Models** Certificate available $90 8 hours	Instructor-Led Workshop **Generative AI with Diffusion Models** Certificate available $500 8 hours

https://www.nvidia.com/en-us/learn/learning-path/generative-ai-llm/

NVIDIA Deep Learning Institute (DLI)



https://learn.nvidia.com/en-us/training/find-training?q=Generative+AI

Generative AI with Diffusion Models

NVIDIA Products Solutions Industries For You Shop Drivers Support Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Self-paced Course

Generative AI with Diffusion Models

Take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines, with applications in creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more.

About Course Objectives Topics Covered Course Outline Stay Informed Contact Us **Continue Learning**

About this Course

Thanks to improvements in computing power and scientific theory, generative AI is more accessible than ever before. Generative AI plays a significant role across industries due to its numerous applications, such as creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more. In this course, learners will take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines.

Learning Objectives

Course Details

Duration: 08:00

Price: $90

Subject: Generative AI/LLM

Language: English

Course Prerequisites:

A basic understanding of Deep Learning Concepts.

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Building RAG Agents with LLMs

NVIDIA. Products Solutions Industries For You

Shop Drivers Support Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Self-paced Course

Building RAG Agents with LLMs

Agents powered by large language models (LLMs) have shown great retrieval capability for using tools, looking at documents, and plan their approaches. This course will show you how to deploy an agent system in practice with the flexibility to scale up your system to meet the demands of users and customers.

About Course Objectives Topics Covered Course Outline Stay Informed Contact Us **Continue Learning**

About this Course

This course is free for a limited time.

The evolution and adoption of large language models (LLMs) have been nothing short of revolutionary, with retrieval-based systems at the forefront of this technological leap. These models are not just tools for automation; they are partners in enhancing productivity, capable of holding informed conversations by interacting with a vast array of tools and documents. This course is designed for those eager to explore the potential of these systems, focusing on practical deployment and the efficient implementation required to manage the considerable demands of both users and deep learning models. As we delve into the intricacies of LLMs, participants will gain insights into advanced orchestration techniques that include internal reasoning, dialog management, and effective tooling strategies.

Course Details

Duration: 08:00

Price: Free

Level: Technical - Intermediate

Subject: Generative AI/LLM

Language: English

Course Prerequisites:
Introductory deep learning knowledge, with comfort

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1

Generative AI Explained

NVIDIA. Products Solutions Industries For You   Shop Drivers Support Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Self-paced Course

Generative AI Explained

In this no-coding course, learn Generative AI concepts and applications, as well as the challenges and opportunities in this exciting field.

About Course Objectives Topics Covered Course Outline Stay Informed Contact Us **Continue Learning**

About this Course

Generative AI describes technologies that are used to generate new content based on a variety of inputs. In recent time, Generative AI involves the use of neural networks to identify patterns and structures within existing data to generate new content. In this course, you will learn Generative AI concepts, applications, as well as the challenges and opportunities in this exciting field.

Learning Objectives

Upon completion, you will have a basic understanding of Generative AI and be able to more effectively use the various tools built on this

Course Details

Duration: 02:00

Price: Free

Level: Technical - Beginner

Subject: Generative AI/LLM

Language: English

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1

Introduction to Transformer-Based Natural Language Processing

NVIDIA. Products Solutions Industries For You | Shop Drivers Support | Min-Yuh Day

Deep Learning Institute Find Training | Self Paced Courses | Instructor-Led Workshops | Educator Programs | Enterprise Solutions | Certification | Resources

Self-paced Course

Introduction to Transformer-Based Natural Language Processing

Learn how Transformers are used as the building blocks of modern large language models (LLMs). You'll then use these models for various NLP tasks, including text classification, named-entity recognition (NER), author attribution, and question answering.

About Course | Objectives | Topics Covered | Course Outline | Stay Informed | Contact Us | **Continue Learning**

About this Course

Large Language Models (LLMs), or Transformers, have revolutionized the field of natural language processing (NLP). Driven by recent advancements, applications of NLP and generative AI have exploded in the past decade. With the proliferation of applications like chatbots and intelligent virtual assistants, organizations are infusing their businesses with more interactive human-machine experiences. Understanding how Transformer-based large language models (LLMs) can be used to manipulate, analyze, and generate text-based data is essential. Modern pre-trained LLMs can encapsulate the nuance, context, and sophistication of language, just as humans do. When fine-tuned and deployed correctly, developers can use these LLMs to build powerful NLP applications that provide natural and seamless human-computer interactions within chatbots, AI voice agents, and more. In this course, you'll learn how Transformers are used as the building blocks of modern large language models (LLMs). You'll then use these models for various NLP

Course Details

Duration: 06:00

Price: $30

Level: Technical - Beginner

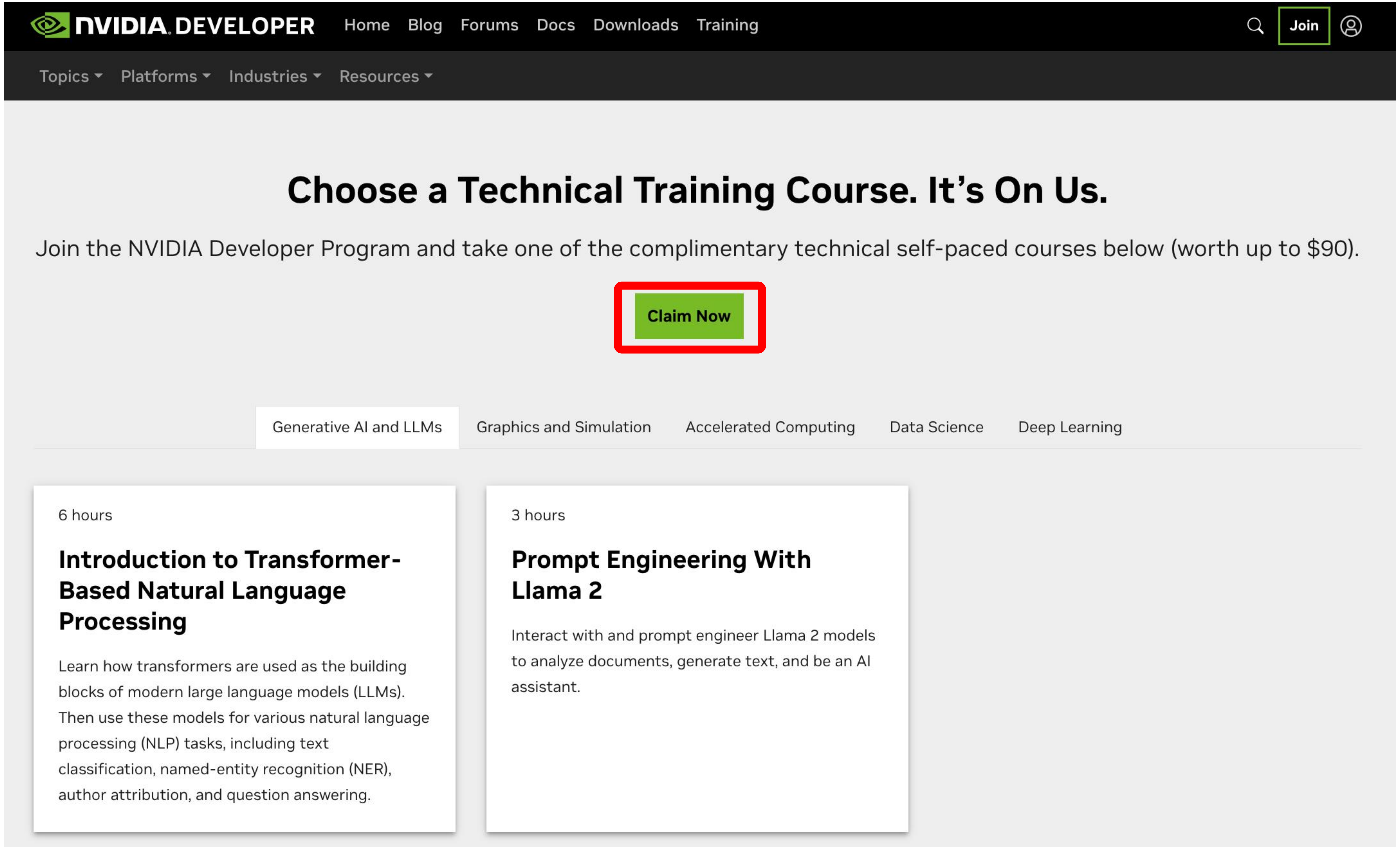
Subject: Generative AI/LLM

Language: English

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-08+V1

Join the NVIDIA Developer Program

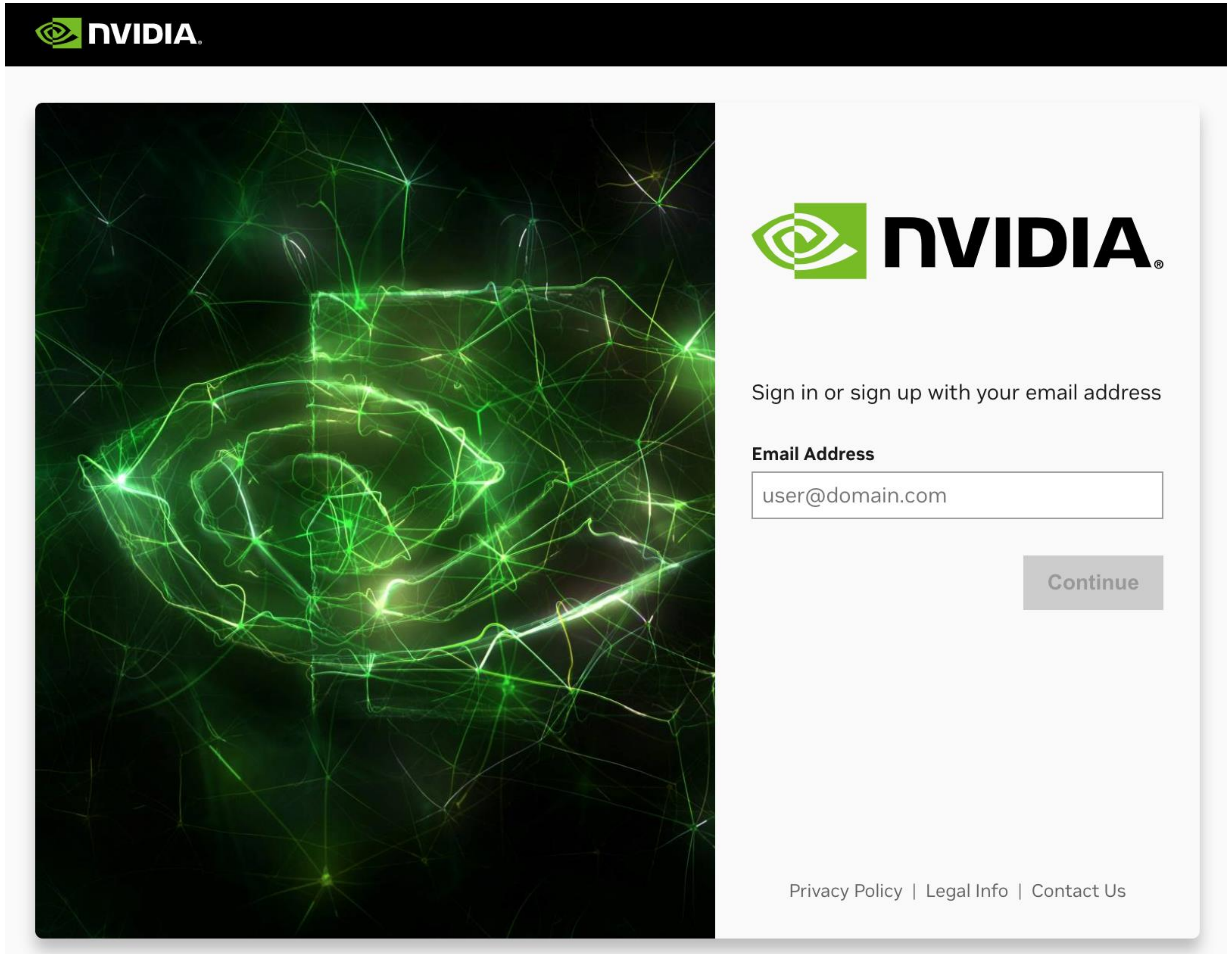
take one of the complimentary technical self-paced courses (worth up to $90)



https://developer.nvidia.com/join-nvidia-developer-program

Join the NVIDIA Developer Program

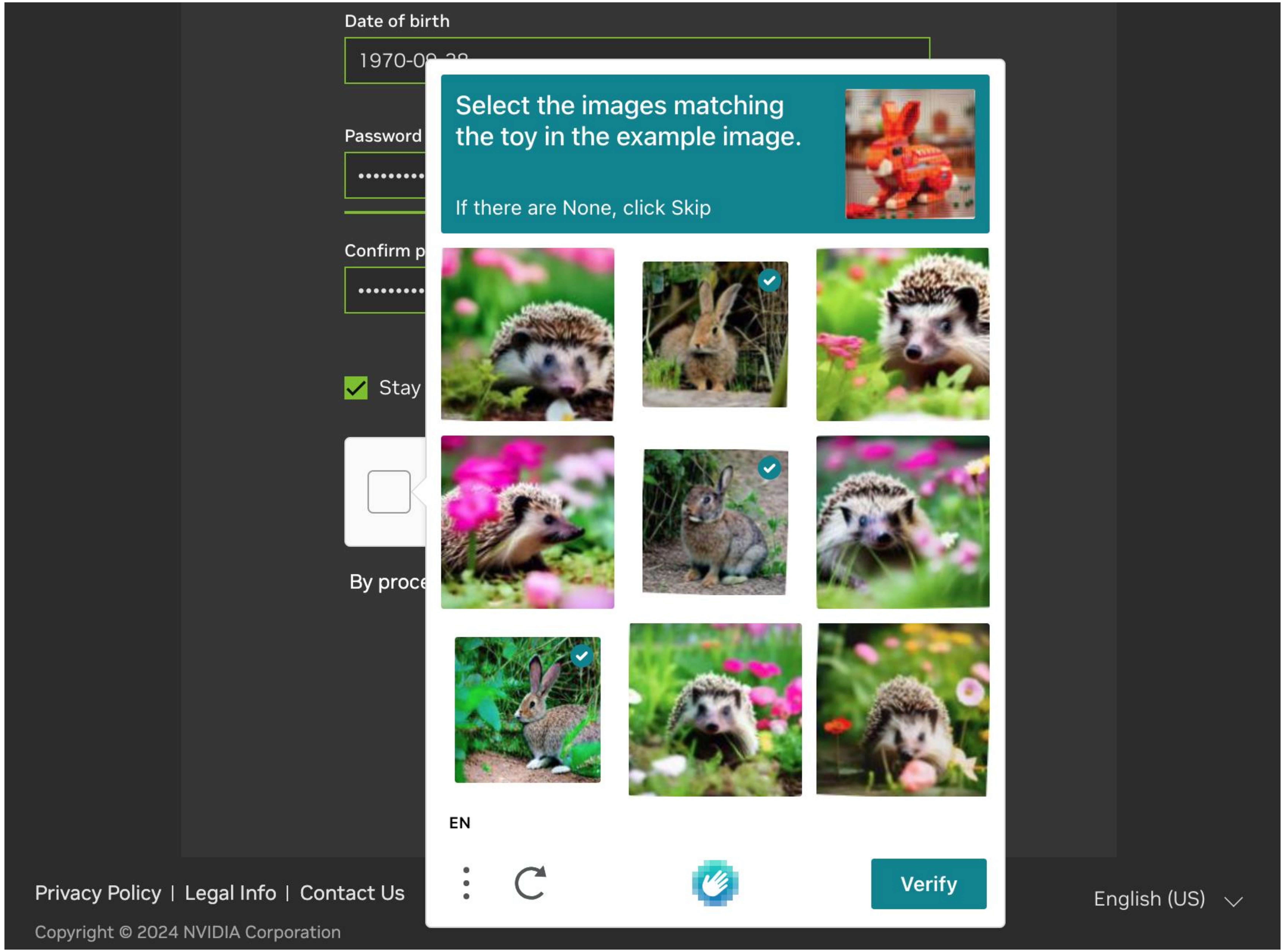
take one of the complimentary technical self-paced courses (worth up to $90)



https://developer.nvidia.com/join-nvidia-developer-program

Join the NVIDIA Developer Program

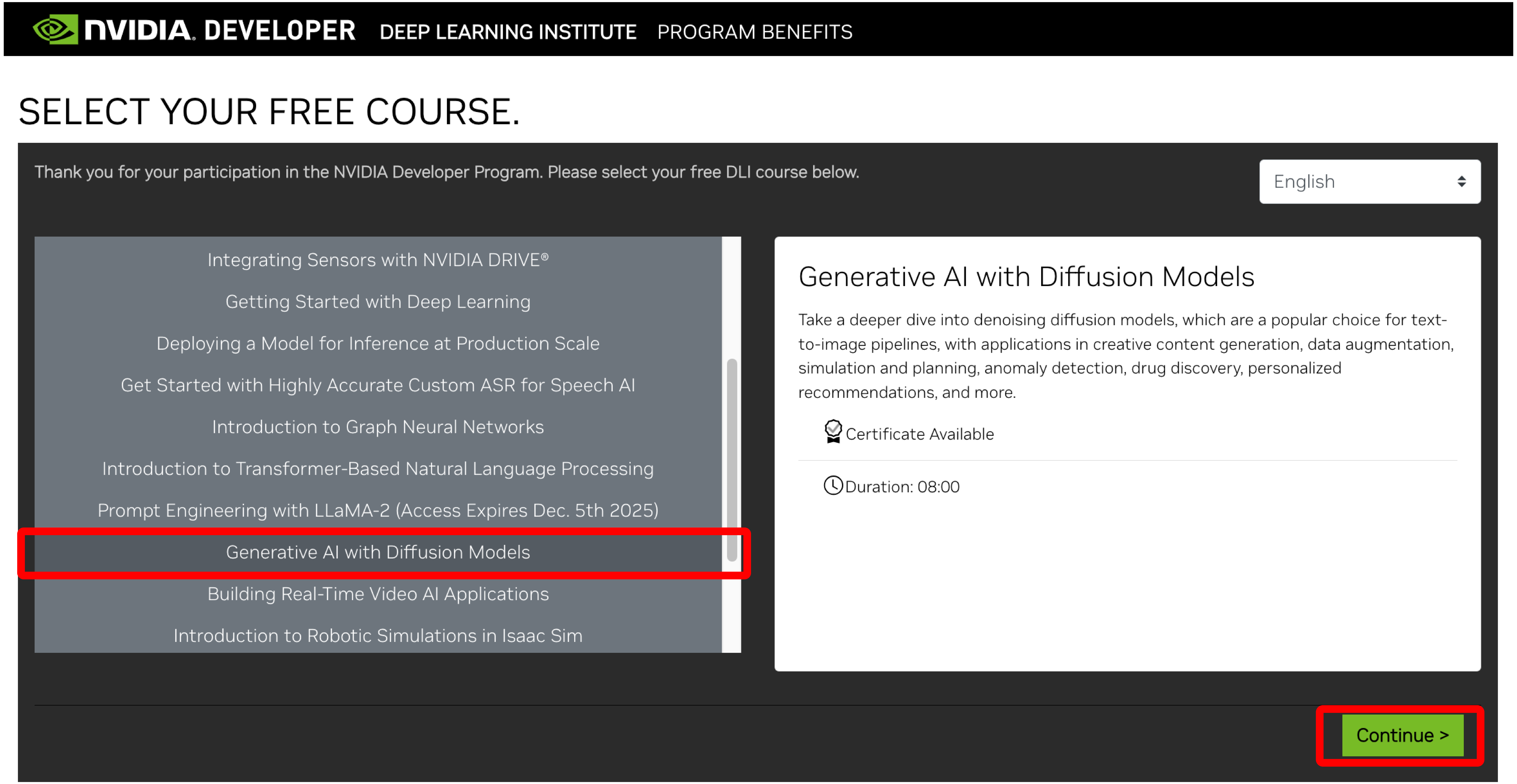
take one of the complimentary technical self-paced courses (worth up to $90)



https://developer.nvidia.com/join-nvidia-developer-program

Join the NVIDIA Developer Program

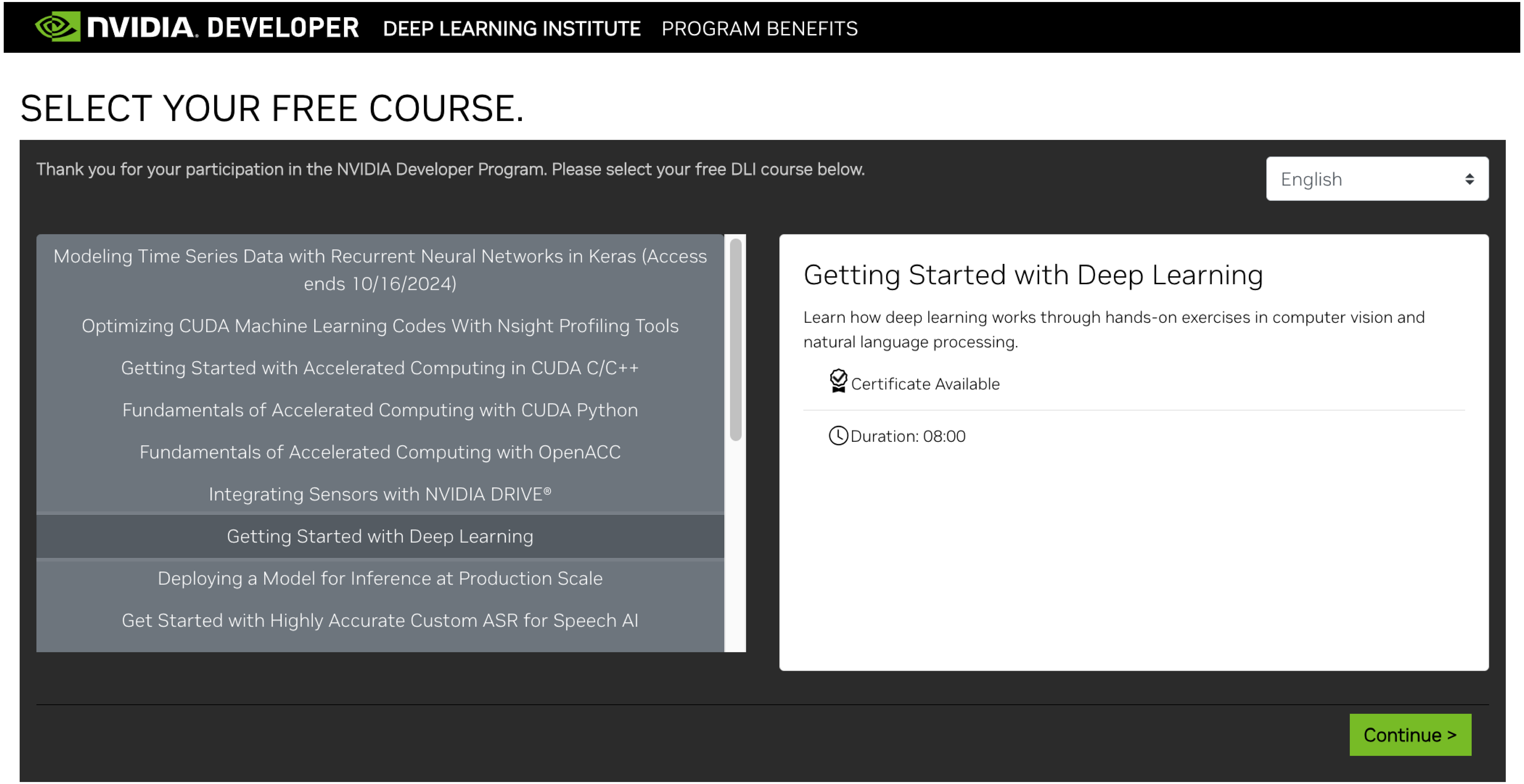
take one of the complimentary technical self-paced courses (worth up to $90)



https://developer.nvidia.com/join-nvidia-developer-program

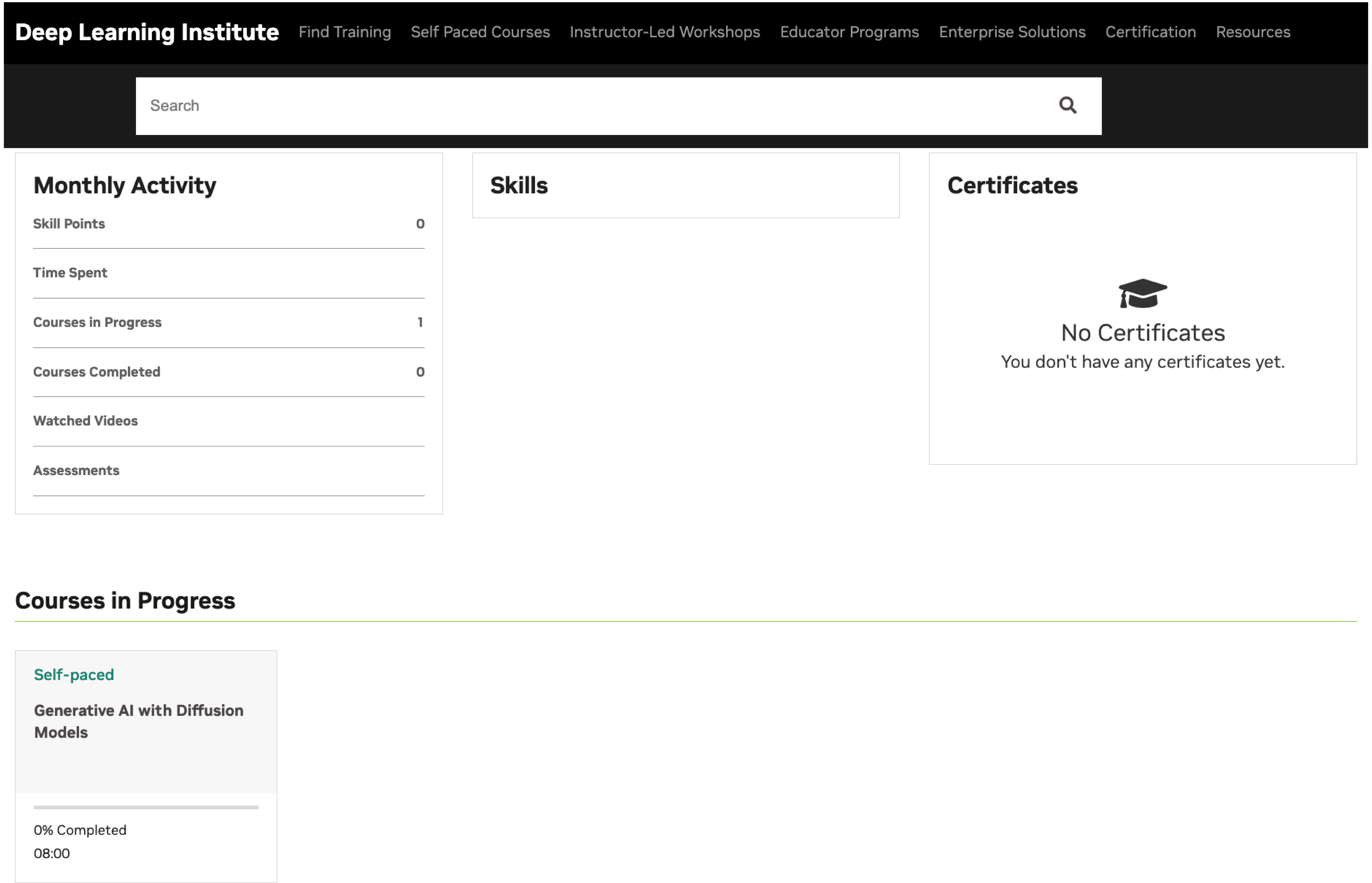
Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to $90)



https://developer.nvidia.com/join-nvidia-developer-program

NVIDIA Deep Learning Institute (DLI)



https://learn.nvidia.com/my-learning

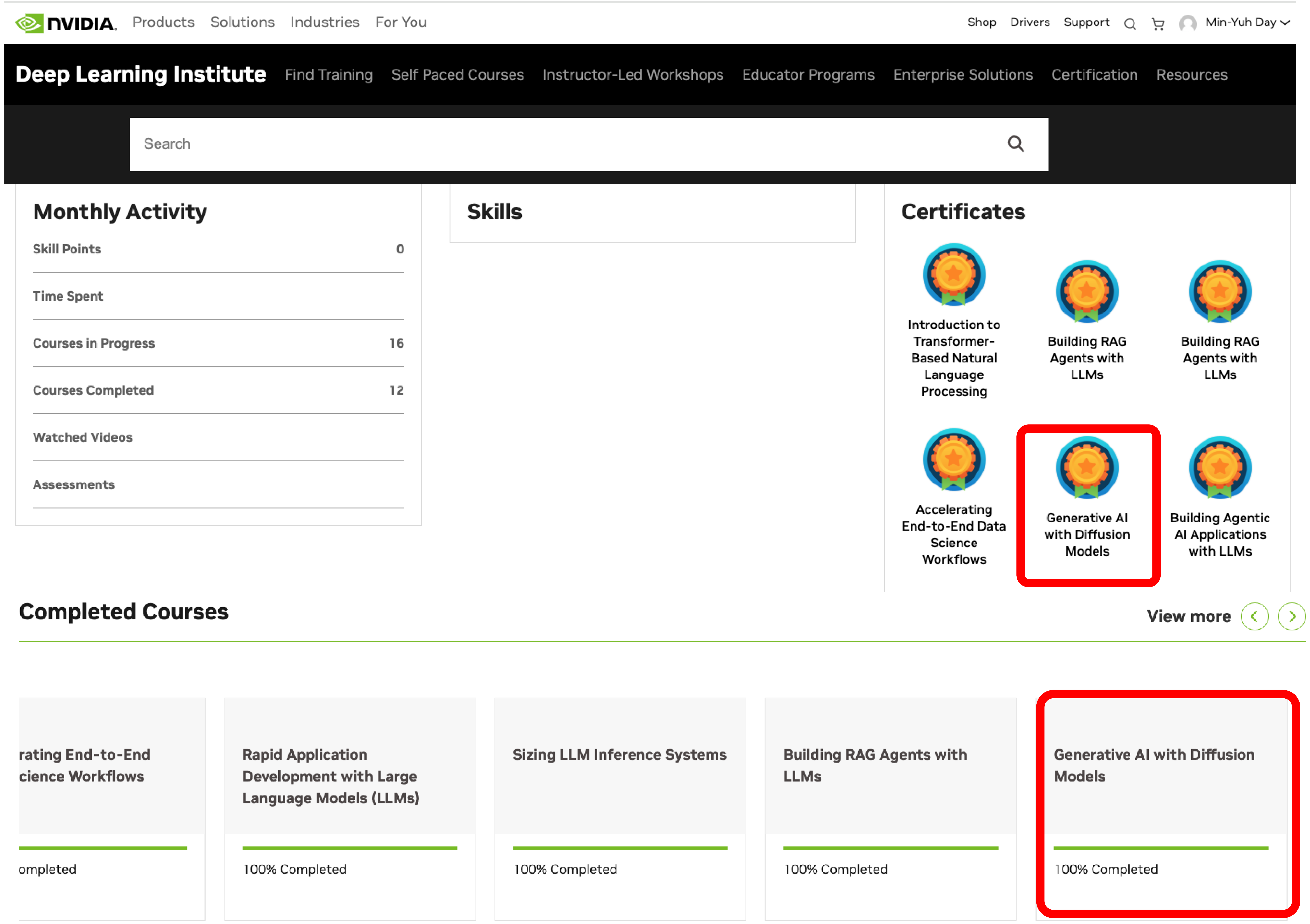
NVIDIA Deep Learning Institute (DLI)

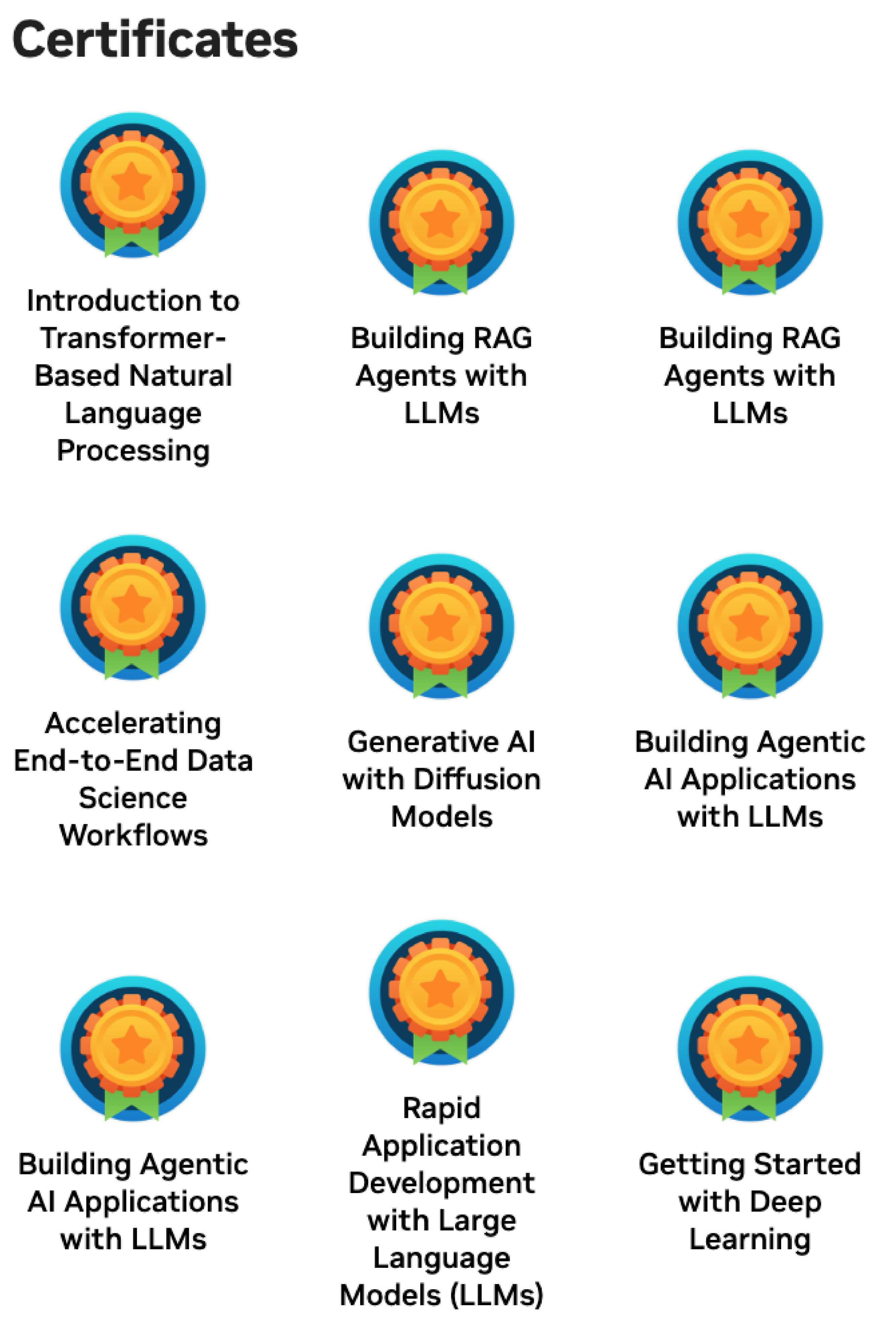
https://learn.nvidia.com/my-learning

NVIDIA Deep Learning Institute (DLI)

https://learn.nvidia.com/my-learning

Deep Learning Institute (DLI)





https://learn.nvidia.com/my-learning

NVIDIA

Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of

Generative AI with Diffusion Models

Greg Estes

Vice President, NVIDIA

Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

https://learn.nvidia.com/certificates?id=q3oo-oBhTQKtyCCote2E-Q

NVIDIA

Certificate of Completion

This certificate is awarded to

Min-Yuh Day

for successfully completing

Building RAG Agents with LLMs

Greg Estes

Vice President, NVIDIA

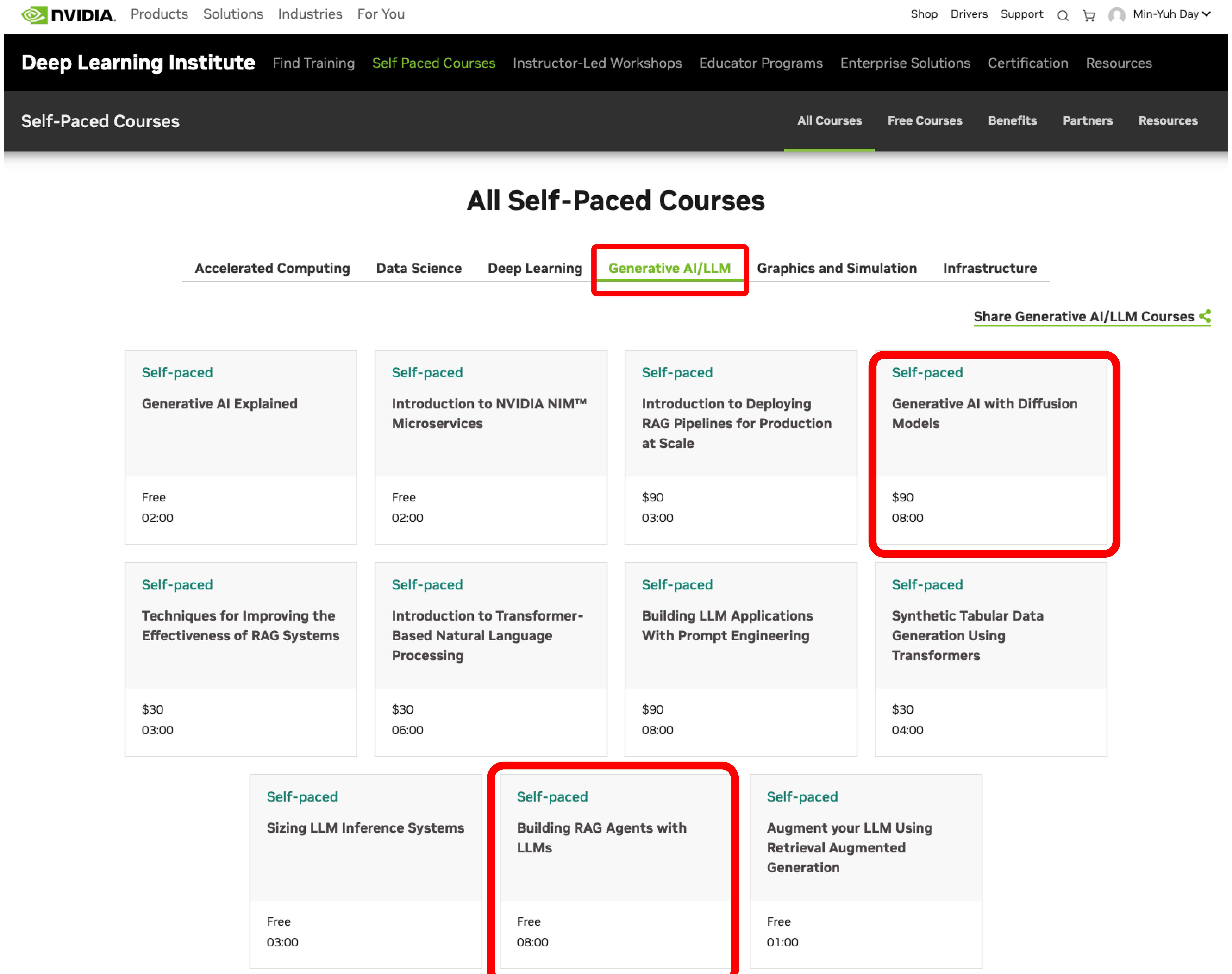
Issue Date: : December 8, 2024

Certification ID: ed-qOCIMQaatzU8SNUNxgw |

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw/courses/course?course_id=course-v1:DLI+S-FX-15+V1

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw

NVIDIA Deep Learning Institute (DLI)



https://learn.nvidia.com/en-us/training/self-paced-courses

Generative AI with Diffusion Models

Products Solutions Industries For You   Shop Drivers Support Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Self-paced Course

Generative AI with Diffusion Models

Take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines, with applications in creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more.

About Course Objectives Topics Covered Course Outline Stay Informed Contact Us   **Continue Learning**

About this Course

Thanks to improvements in computing power and scientific theory, generative AI is more accessible than ever before. Generative AI plays a significant role across industries due to its numerous applications, such as creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more. In this course, learners will take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines.

Learning Objectives

Course Details

Duration: 08:00

Price: $90

Subject: Generative AI/LLM

Language: English

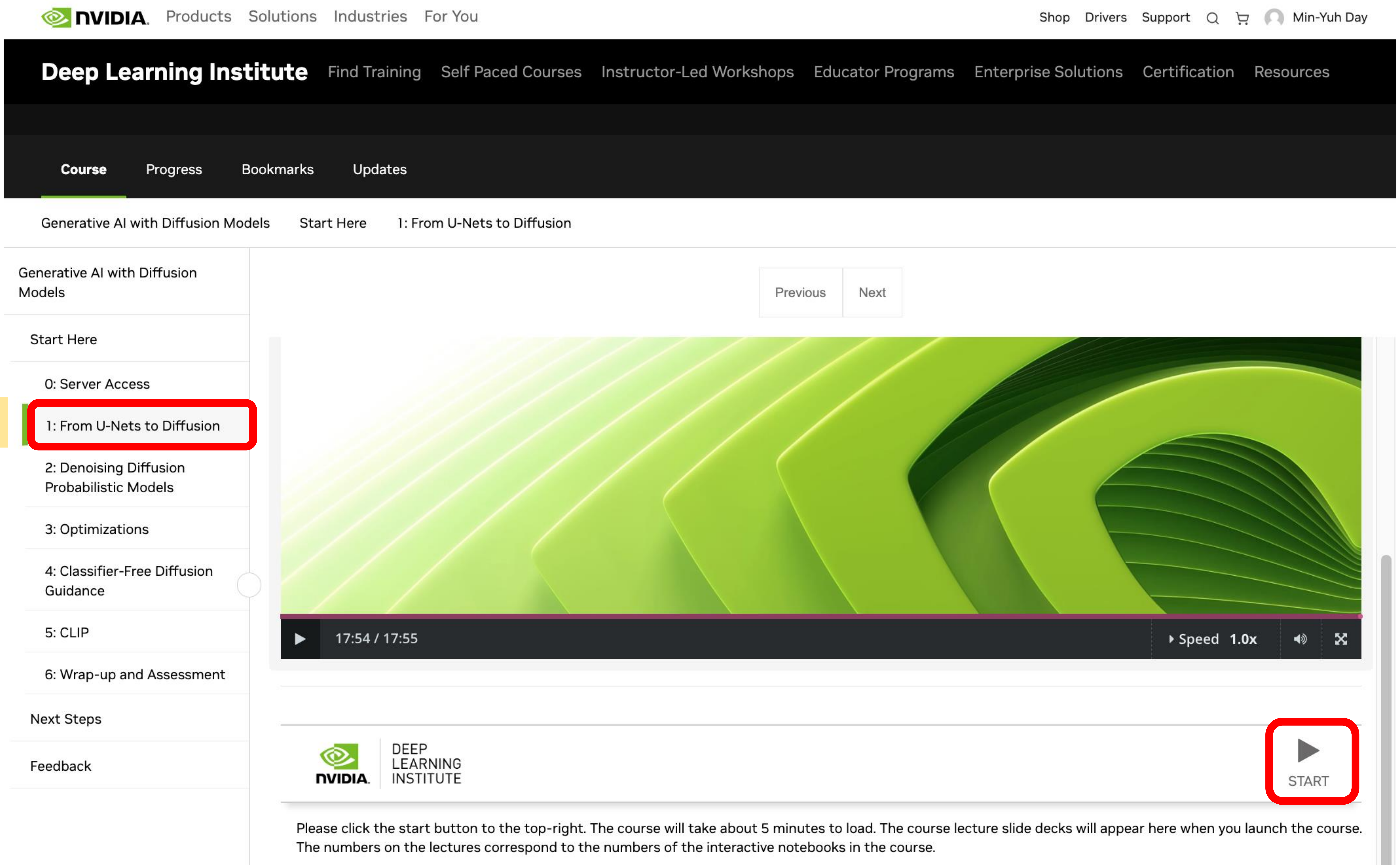
Course Prerequisites:

A basic understanding of Deep Learning Concepts.

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

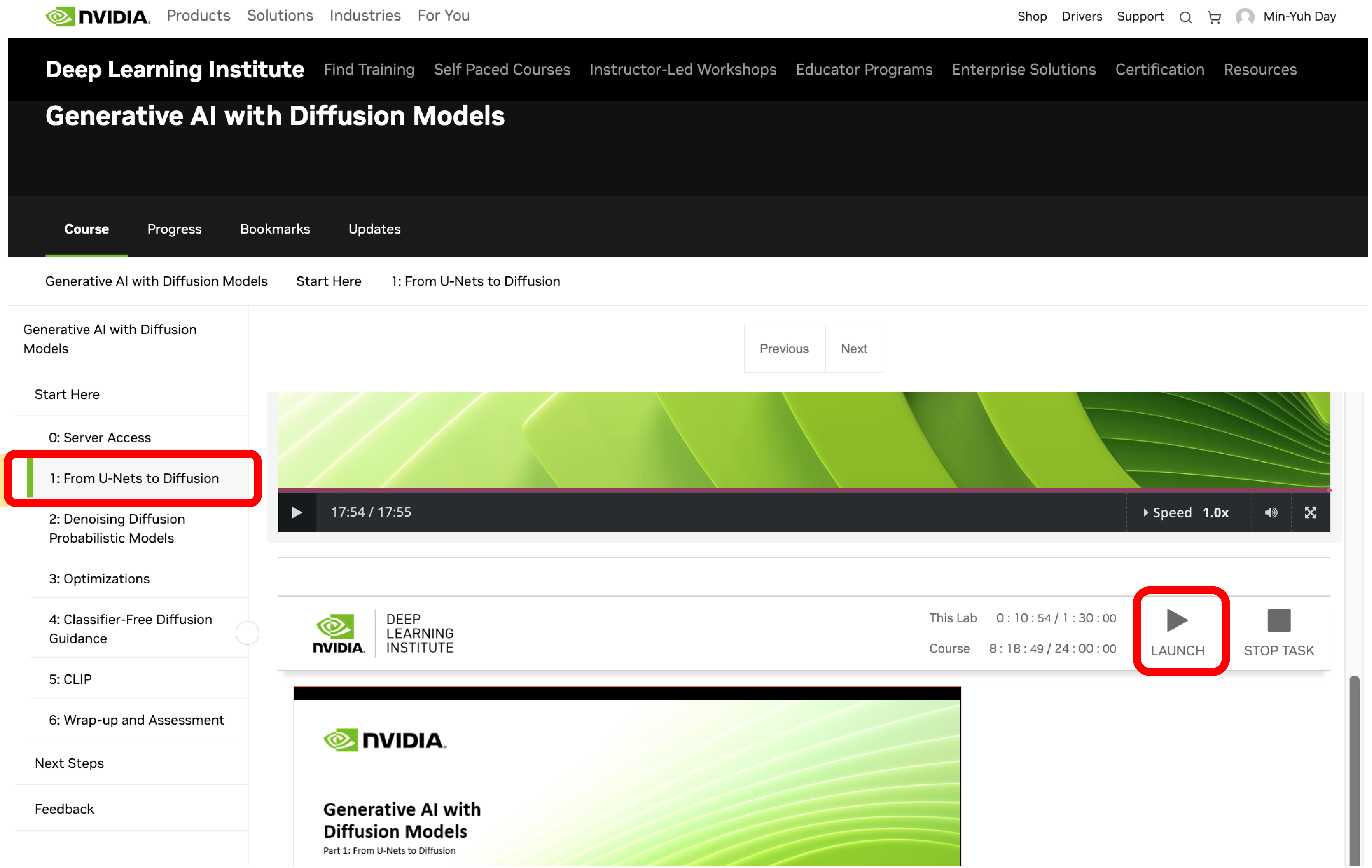
1: From U-Nets to Diffusion



https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

1: From U-Nets to Diffusion

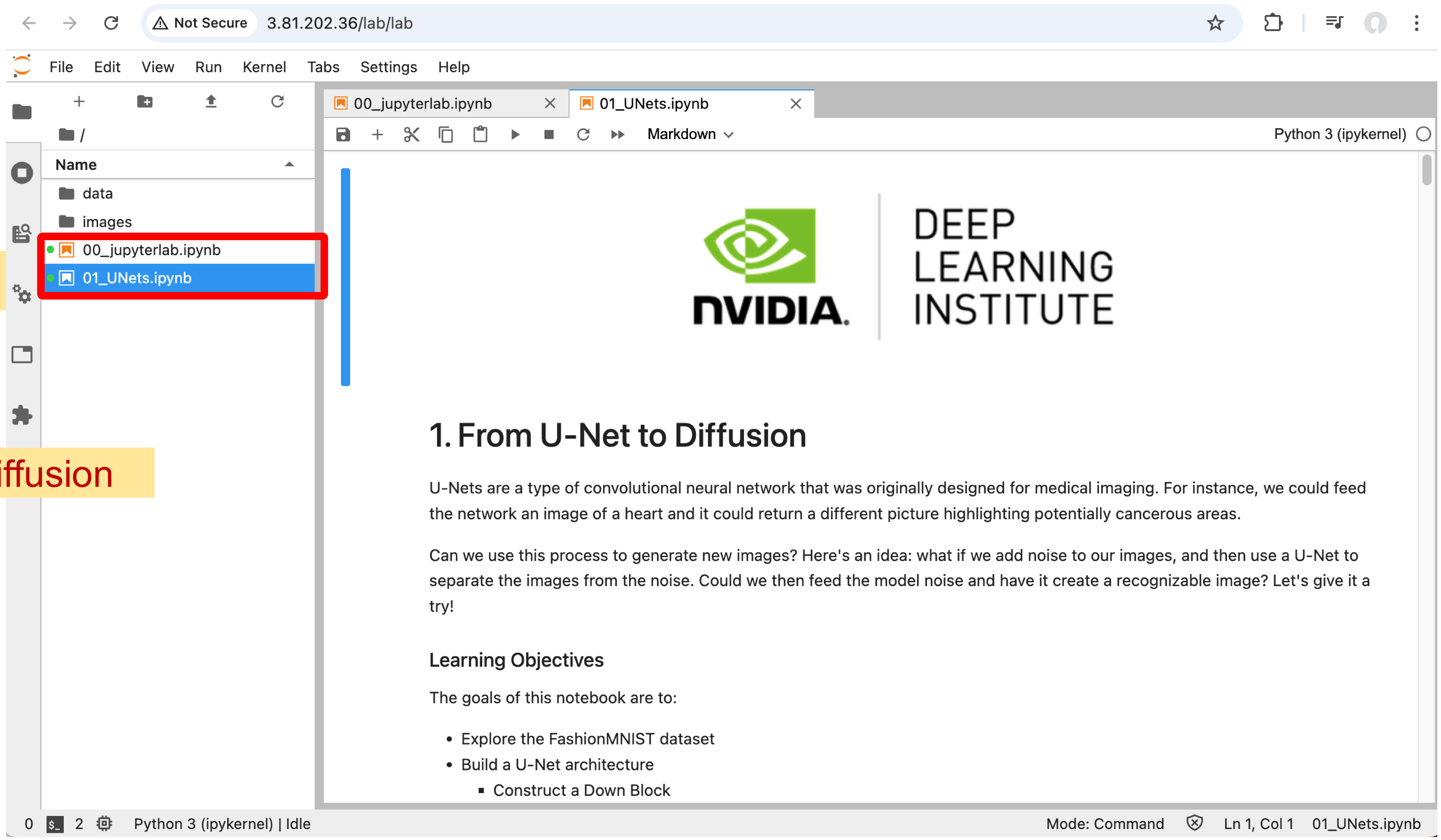


https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

01_UNets.ipynb

1: From U-Nets to Diffusion



https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

Part 1: From U-Nets to Diffusion

Agenda

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- Part 3: Optimizations
- Part 4: Classifier Free Diffusion
- Part 5: CLIP
- Part 6: Wrap-up & Assessment

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Prerequisites

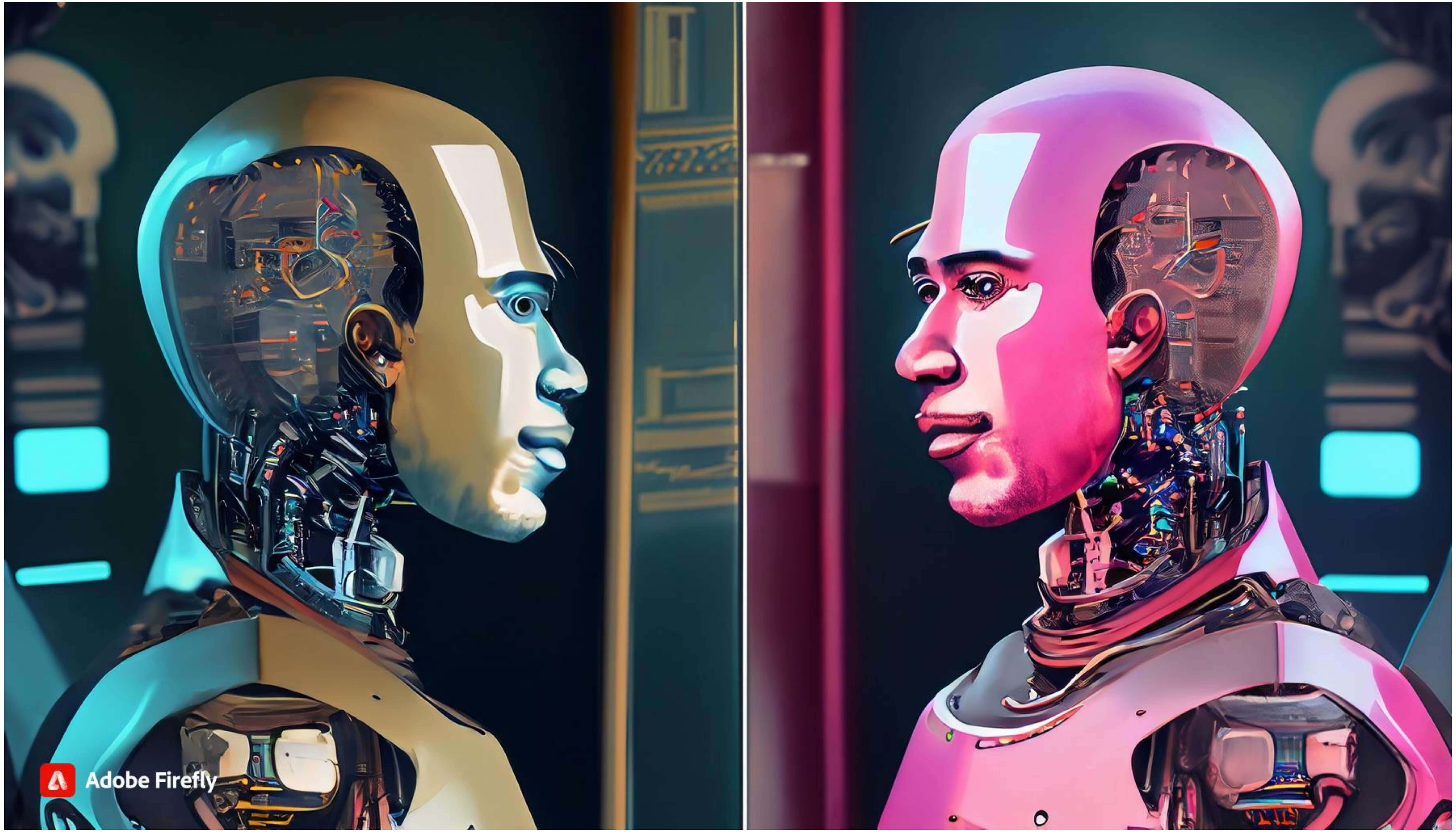
- Basic familiarity with convolutional neural networks (CNNs)
- Basic familiarity with a deep learning framework such as:
  - PyTorch
  - TensorFlow

A Brief History of Generative AI

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

The Imitation Game

A.K.A The Turing Test

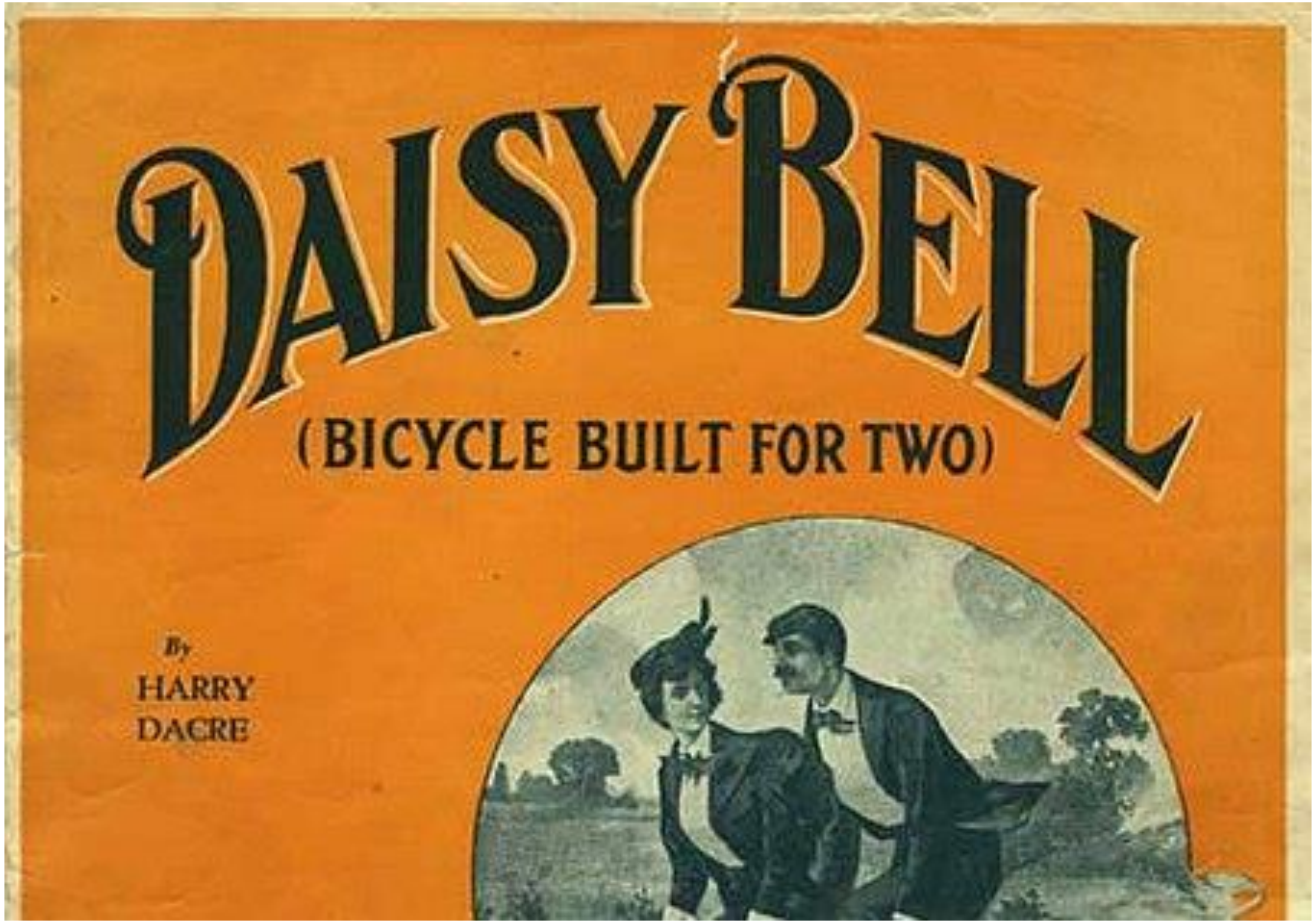


A robot looks in a mirror and the reflection is human, cyberpunk

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

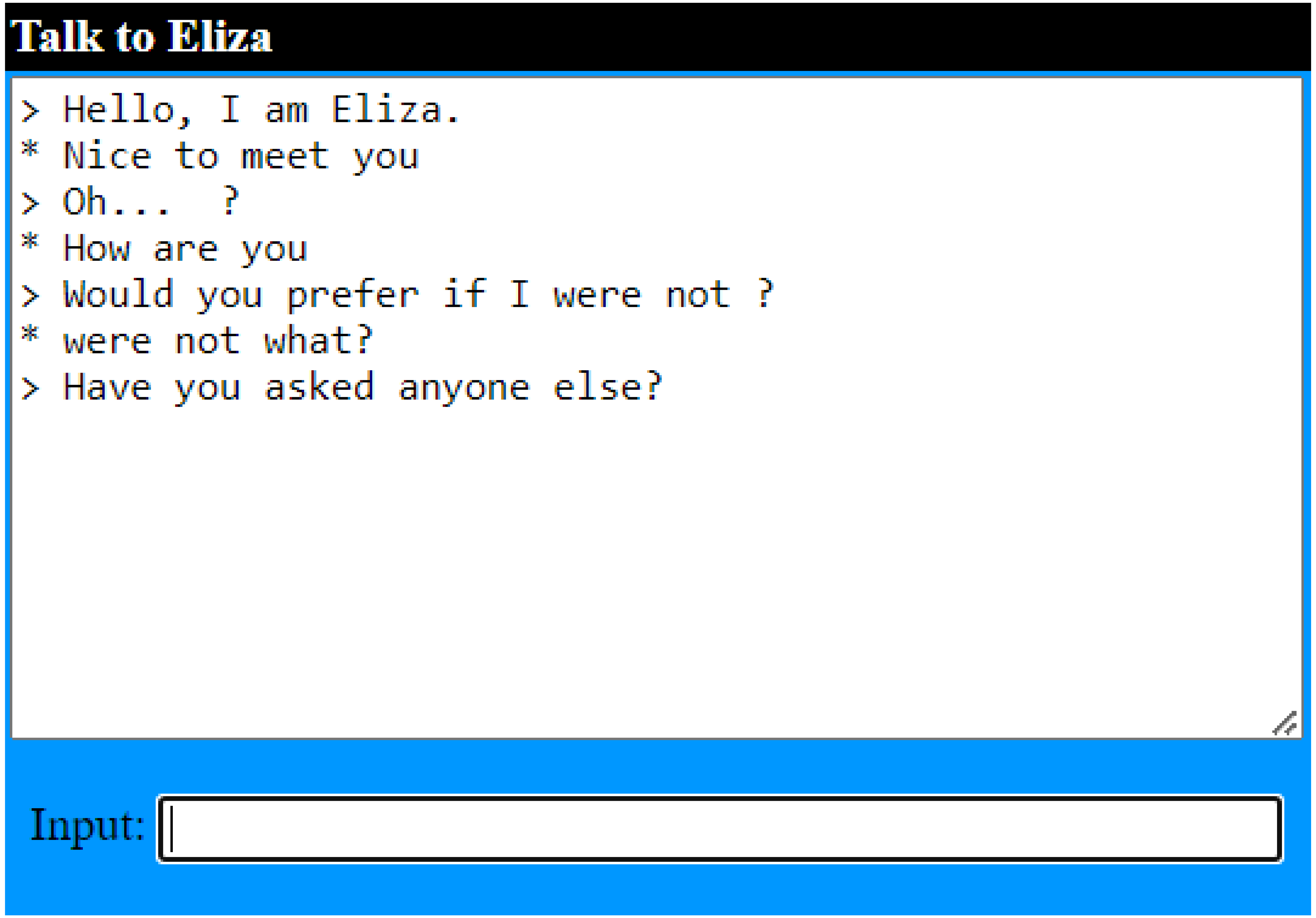
IBM 704

The First Singing Computer



Eliza

The First Gen AI Chatbot?



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI of the 70’s, 80’s and 90s?

- Electronic music
- Video games graphics
- Video game AI
- Computer animation
- Instant messaging chatbots

An 80’s arcade with lots of machines

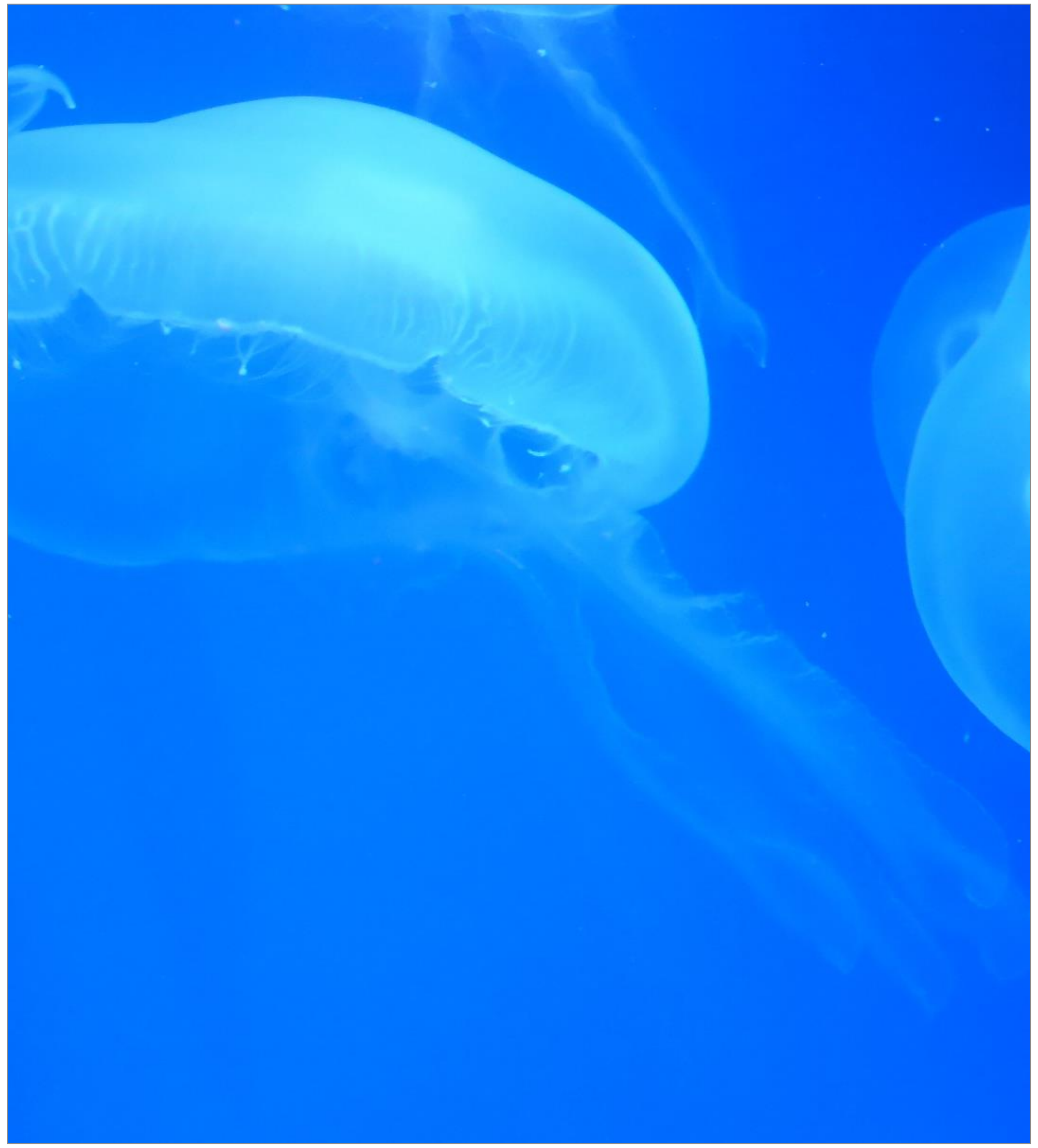
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

The Rise of Neural Networks

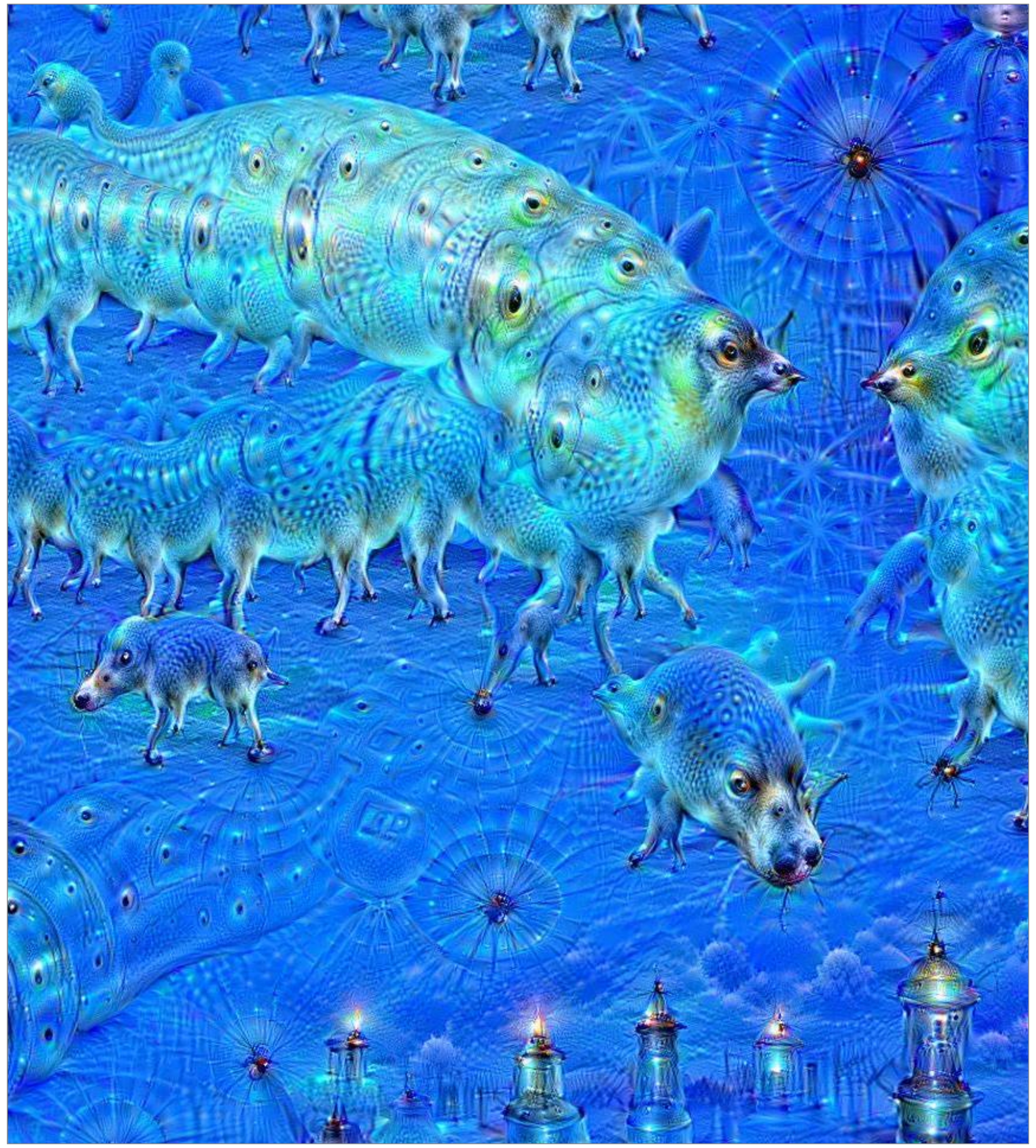
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Deep Dreaming

Images by Martin Thoma



Original



10 Iterations

50 Iterations

GANs

Generative Adversarial Networks

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

GANs

Generative Adversarial Networks

GANs

Generative Adversarial Networks

GANs

Generative Adversarial Networks

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

GANs

Generative Adversarial Networks

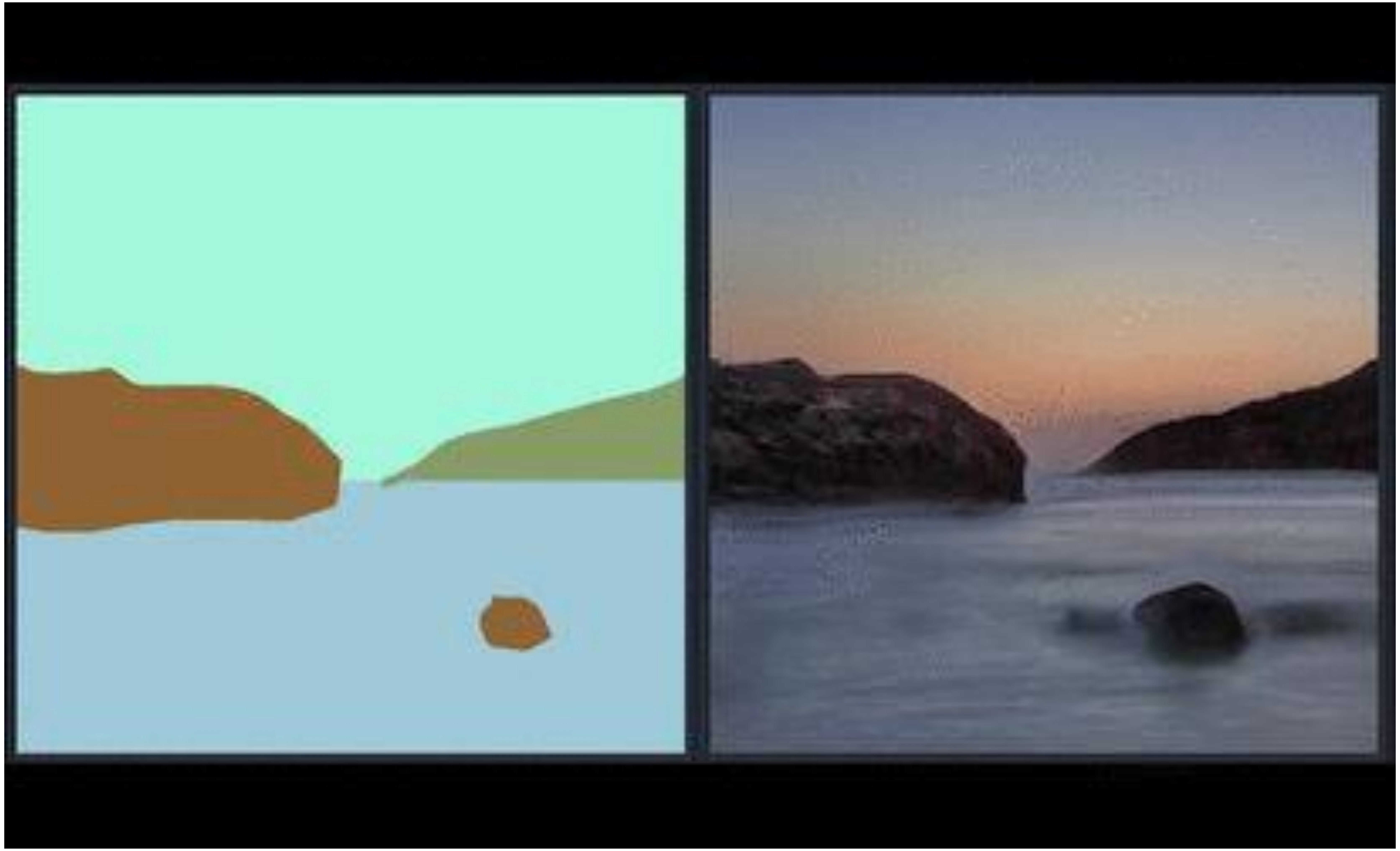
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Image Segmentation

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Image Segmentation + GANs

NVIDIA Spade



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

U-Nets

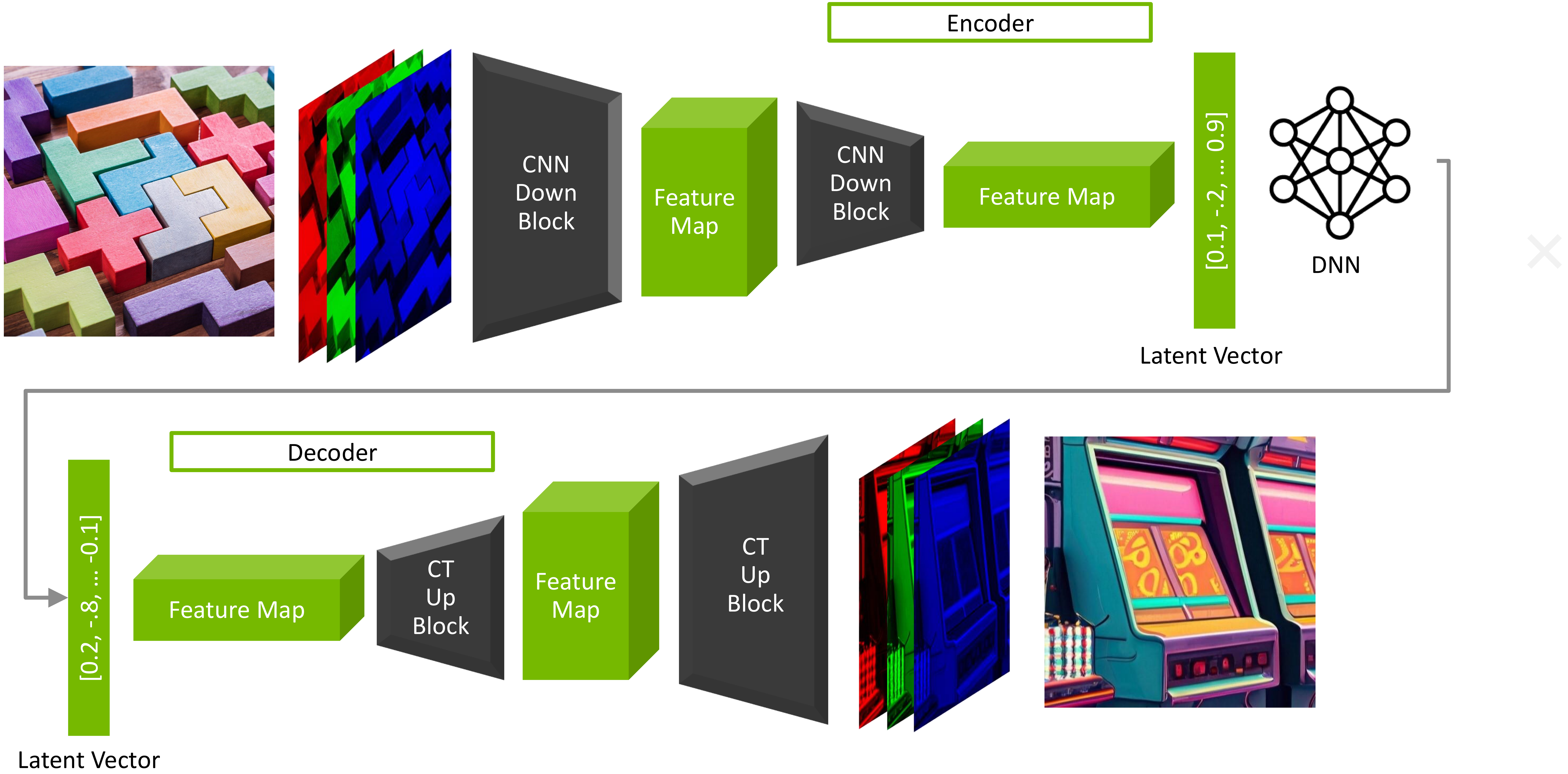
GANs

Generative Adversarial Networks

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

~~GANs~~ U-Nets

The U shaped Autoencoder



U-Nets

The U shaped Autoencoder

Transposed Convolution

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
0	1	0
1	0	1

Output

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1 • .25	0 • .25	1
0 • .25	1 • .25	0
1	0	1

Output

.5	

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0 • .25	1 • .25
0	1 • .25	0 • .25
1	0	1

Output

.5	.5

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
0 • .25	1 • .25	0
1 • .25	0 • .25	1

Output

.5	.5
.5	

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
0	1 • .25	0 • .25
1	0 • .25	1 • .25

Output

.5	.5
.5	.5

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

1	0	1
0	1	0
1	0	1

Output

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1 • .25	0 • .25	0	0	1
0 • .25	0 • .25	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25			

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0 • .25	0 • .25	0	1
0	0 • .25	0 • .25	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25	0		

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25	0	0	.25
0	.25	.25	0
0	.25	.25	0
.25	0	0	.25

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Stride

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

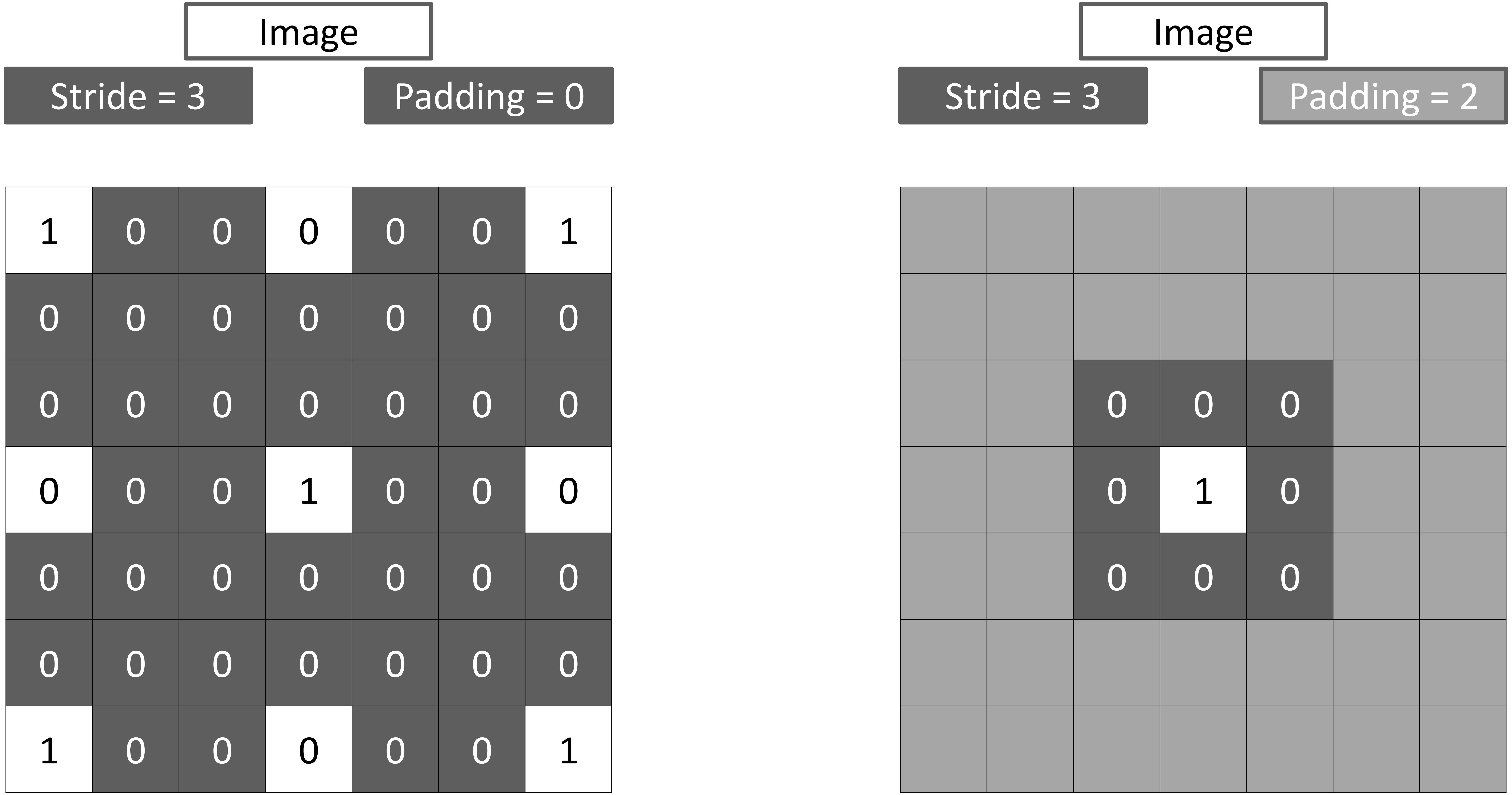
Image

Stride = 3

1	0	0	0	0	0	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	0	0	0	0	0	1

Transposed Convolution

Padding



Transposed Convolution

Padding

Image

Stride = 3

Padding = 0

1	0	0	0	0	0	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	0	0	0	0	0	1

Image

Stride = 3

Padding = 2

0	0	0
0	1	0
0	0	0

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Transposed Convolution

Out Padding

Image

Out Padding = 0

1	0	1
0	1	0
1	0	1

Image

Out Padding = 1

1	0	1	0
0	1	0	0
1	0	1	0
0	0	0	0

Image

Out Padding = 2

1	0	1	0	0
0	1	0	0	0
1	0	1	0	0
0	0	0	0	0
0	0	0	0	0

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Image Resizing

Upsampling

64 px

64 px

128 px

128 px

192 px

192 px

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

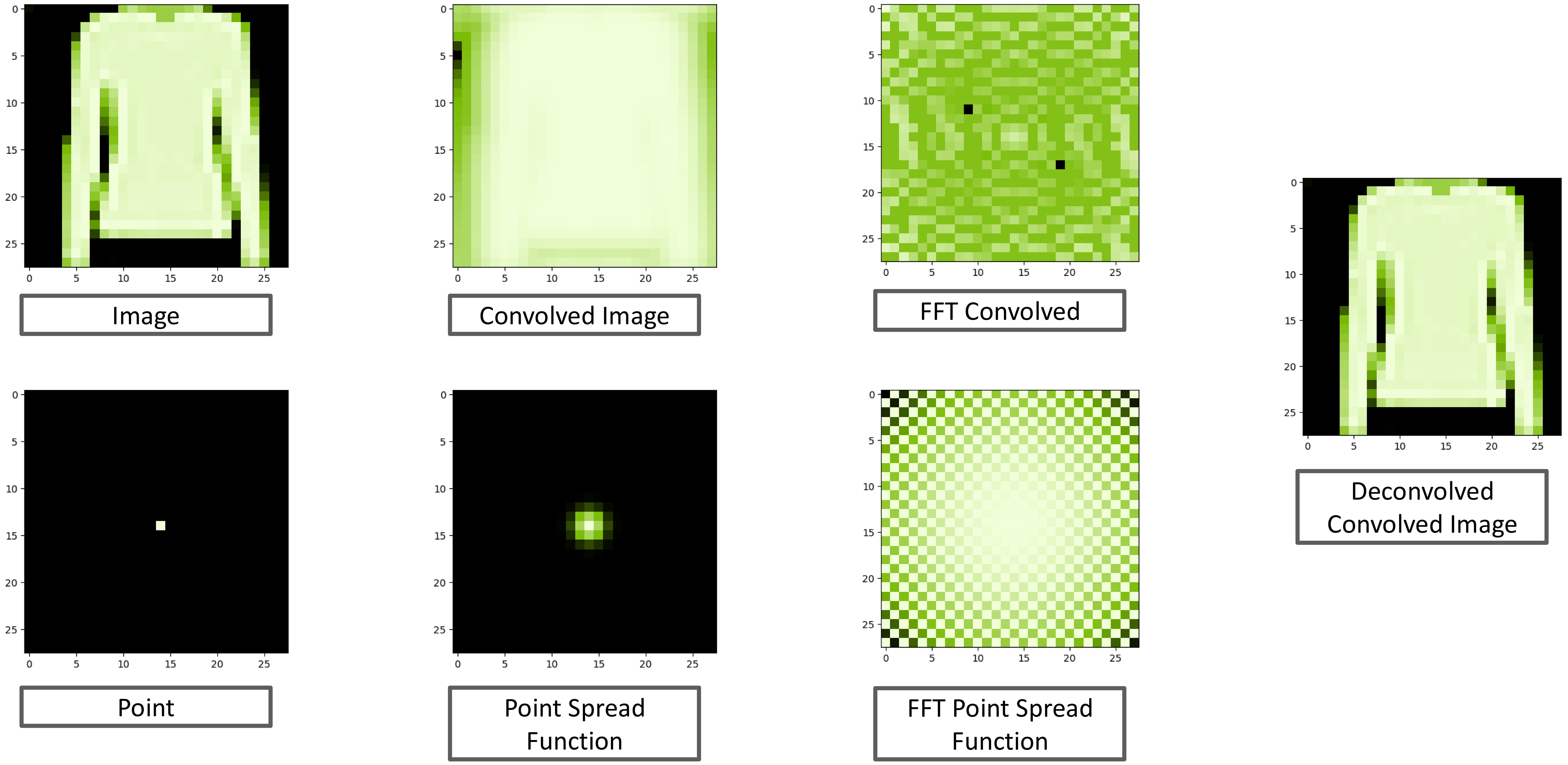
Image Resizing

Upsampling

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Deconvolution?

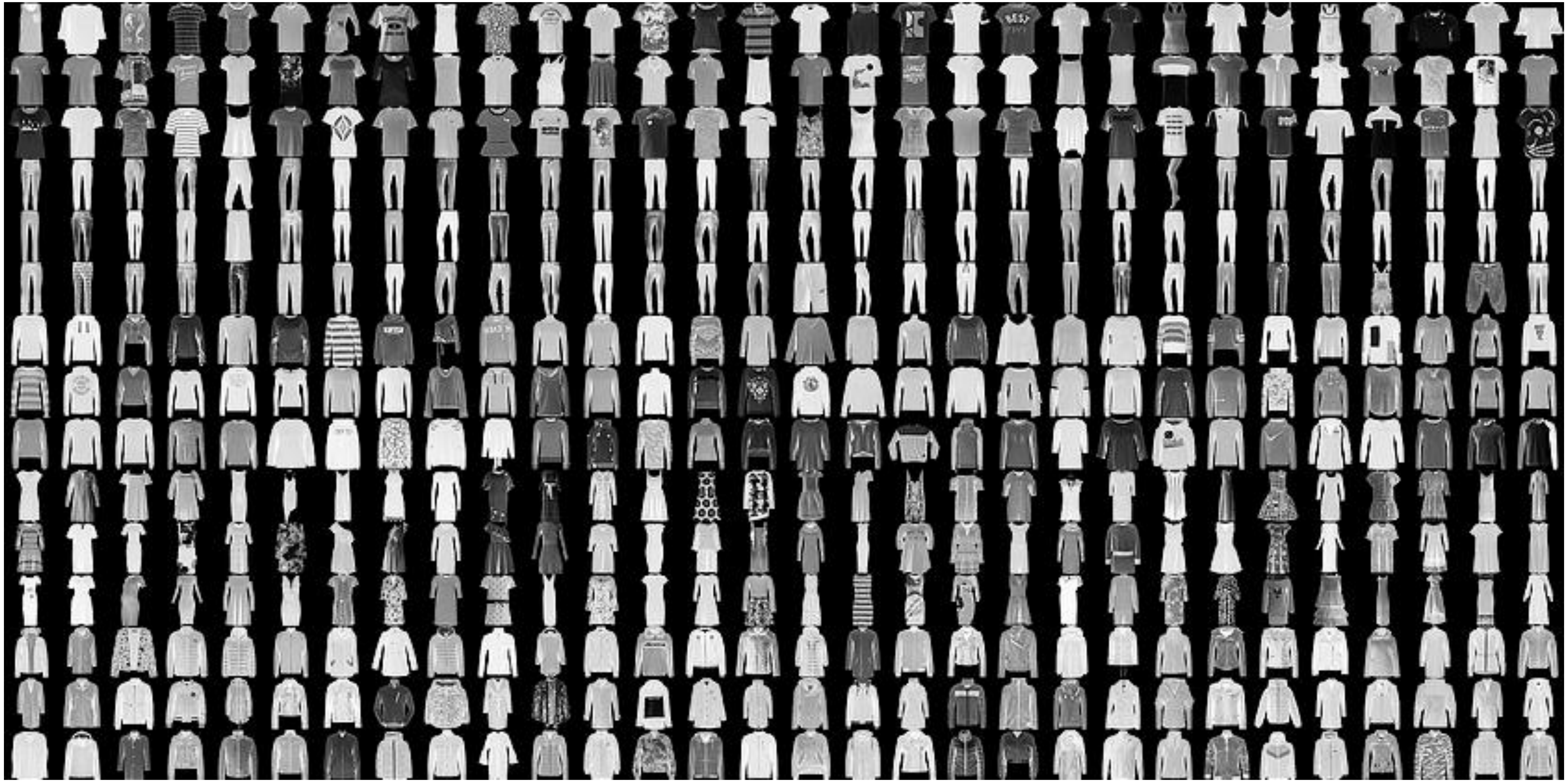
Same as Transposed Convolution?



Lab

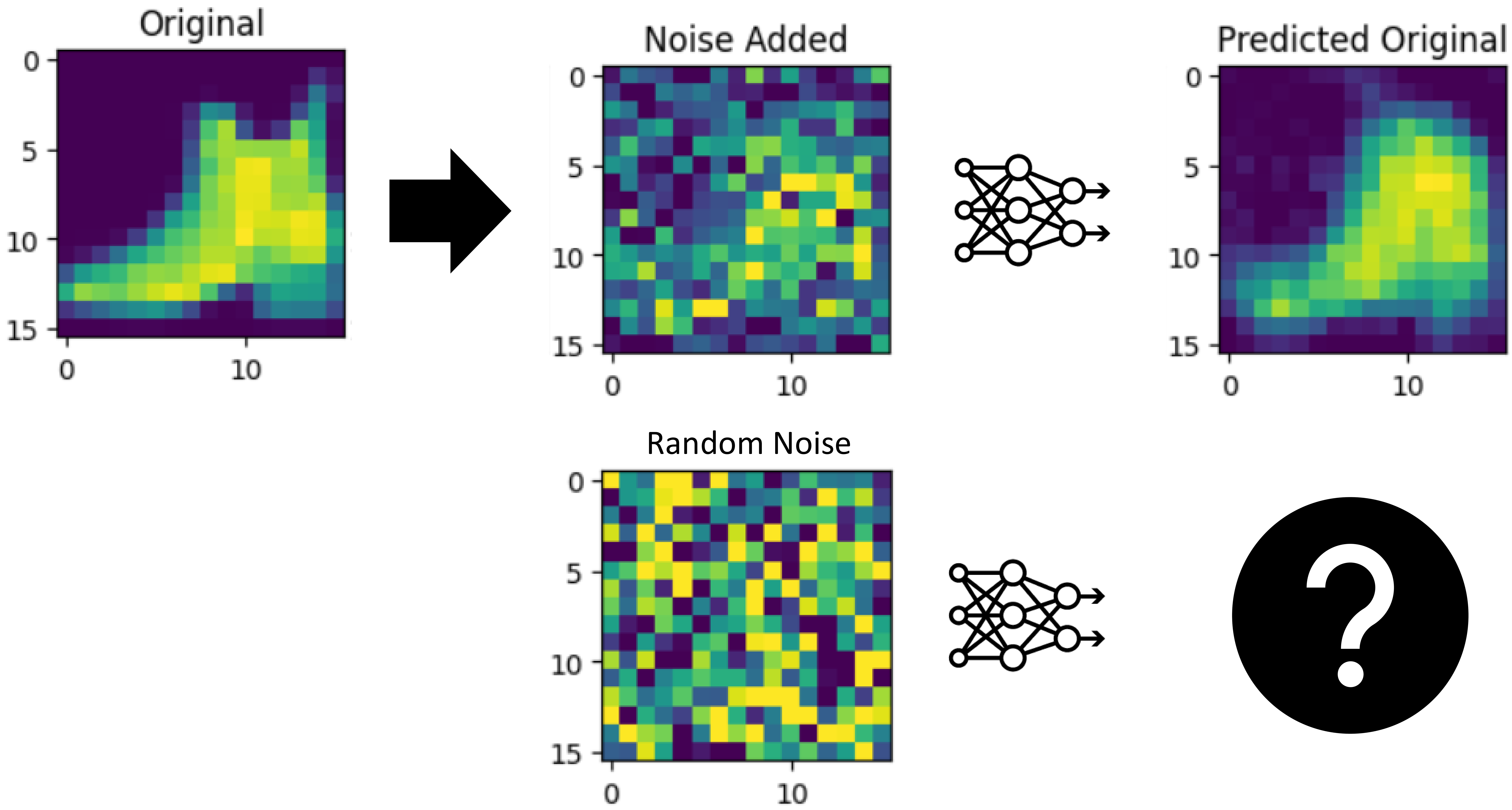
FashionMNIST

Convolutional Neural Network “Hello World”



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Hypothesis: Generate an image from Noise



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

The Experiment

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Let's get started!

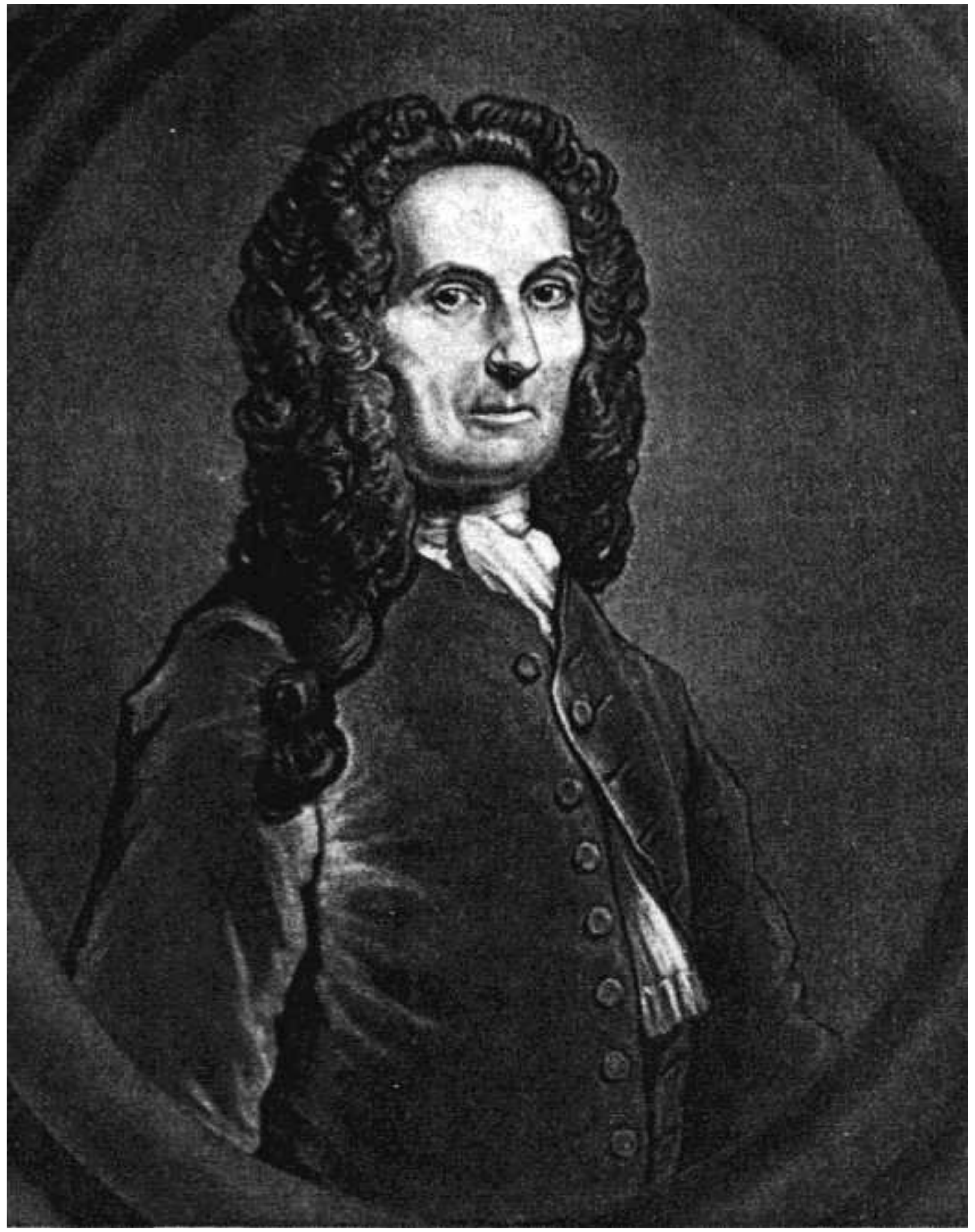
1: From U-Nets to Diffusion

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Appendix:
The Normal Distribution

De Moivre

From Coin Flips to Bells



$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$p = \frac{1}{2}$$

$$n = 4$$

$$\Pr(X = k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

$$k = 2$$

$$\Pr(X = 2) = \frac{4!}{2!\,(4-2)!} \left(\frac{1}{2}\right)^2 (1 - \frac{1}{2})^{4-2}$$

$$\Pr(X = 2) = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} \left(\frac{1}{4}\right) \left(\frac{1}{4}\right)$$

$$\Pr(X = 2) = \frac{6}{16}$$

A weighted coin flipping through the air like a cartoon

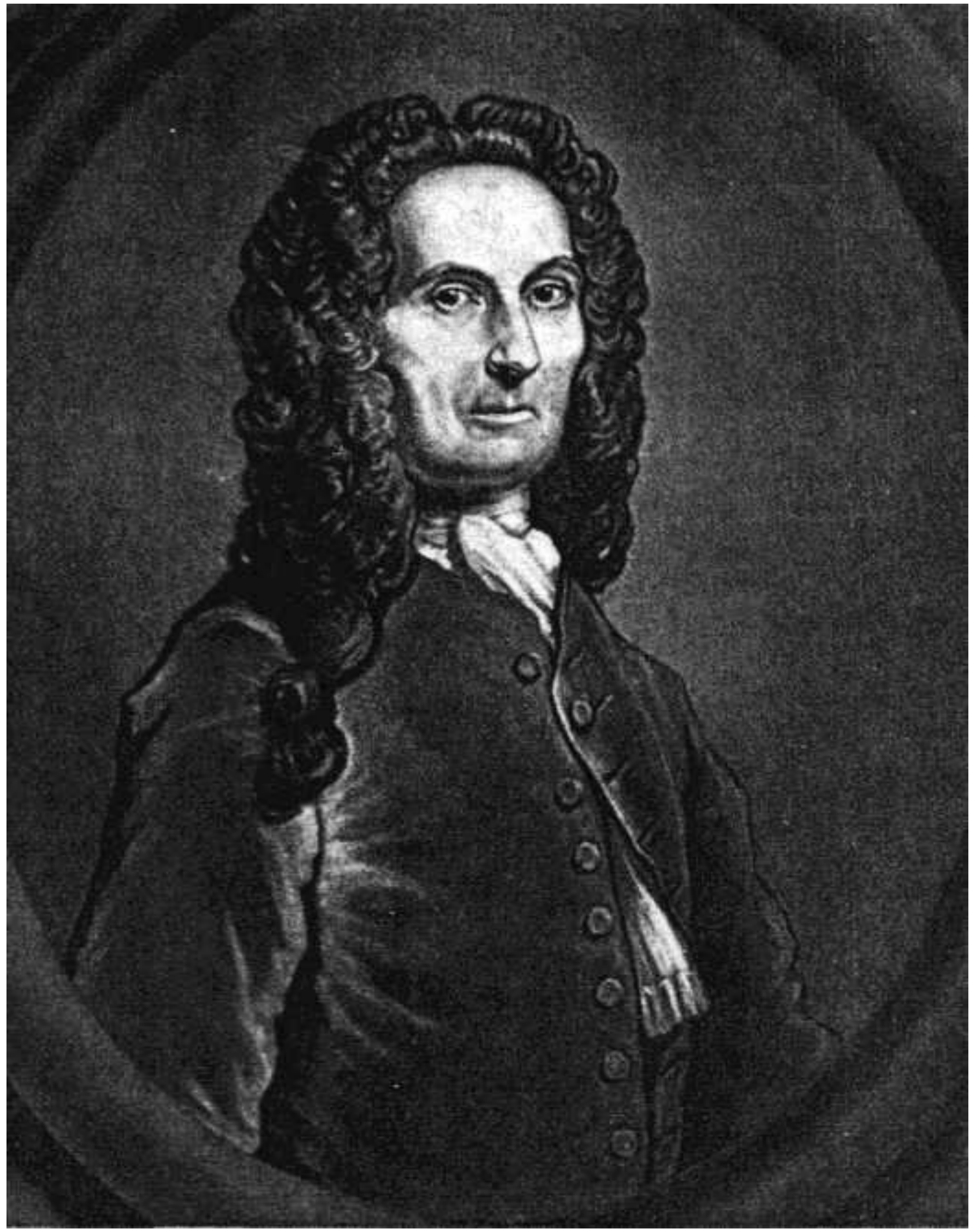
De Moivre

From Coin Flips to Bells

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

De Moivre

From Coin Flips to Bells



$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\Pr(X = k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

$$n! \approx \sqrt{2\pi} \left(\frac{n}{e}\right)^n$$

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}$$

$$N(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

A weighted coin flipping through the air like a cartoon

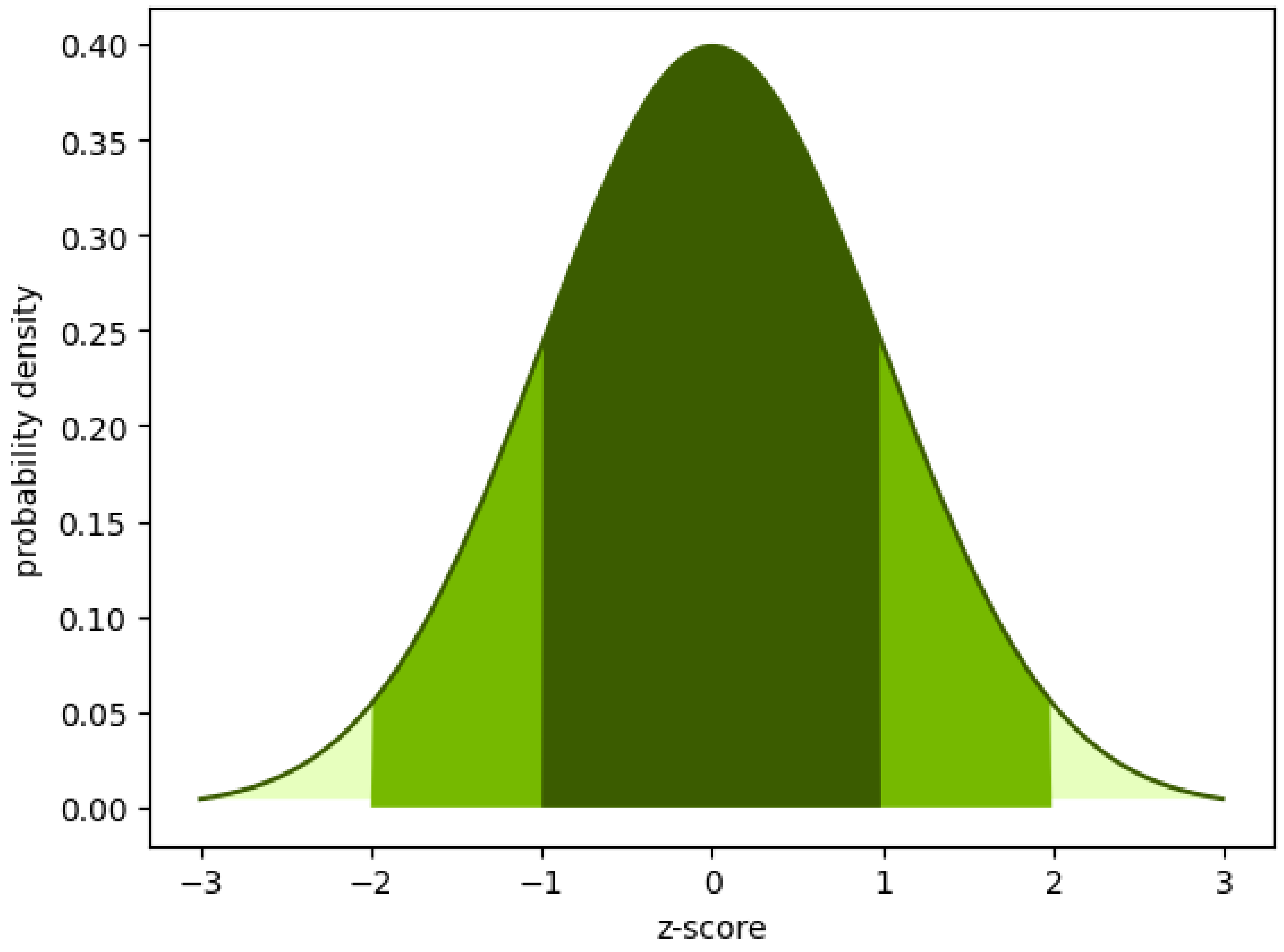
De Moivre

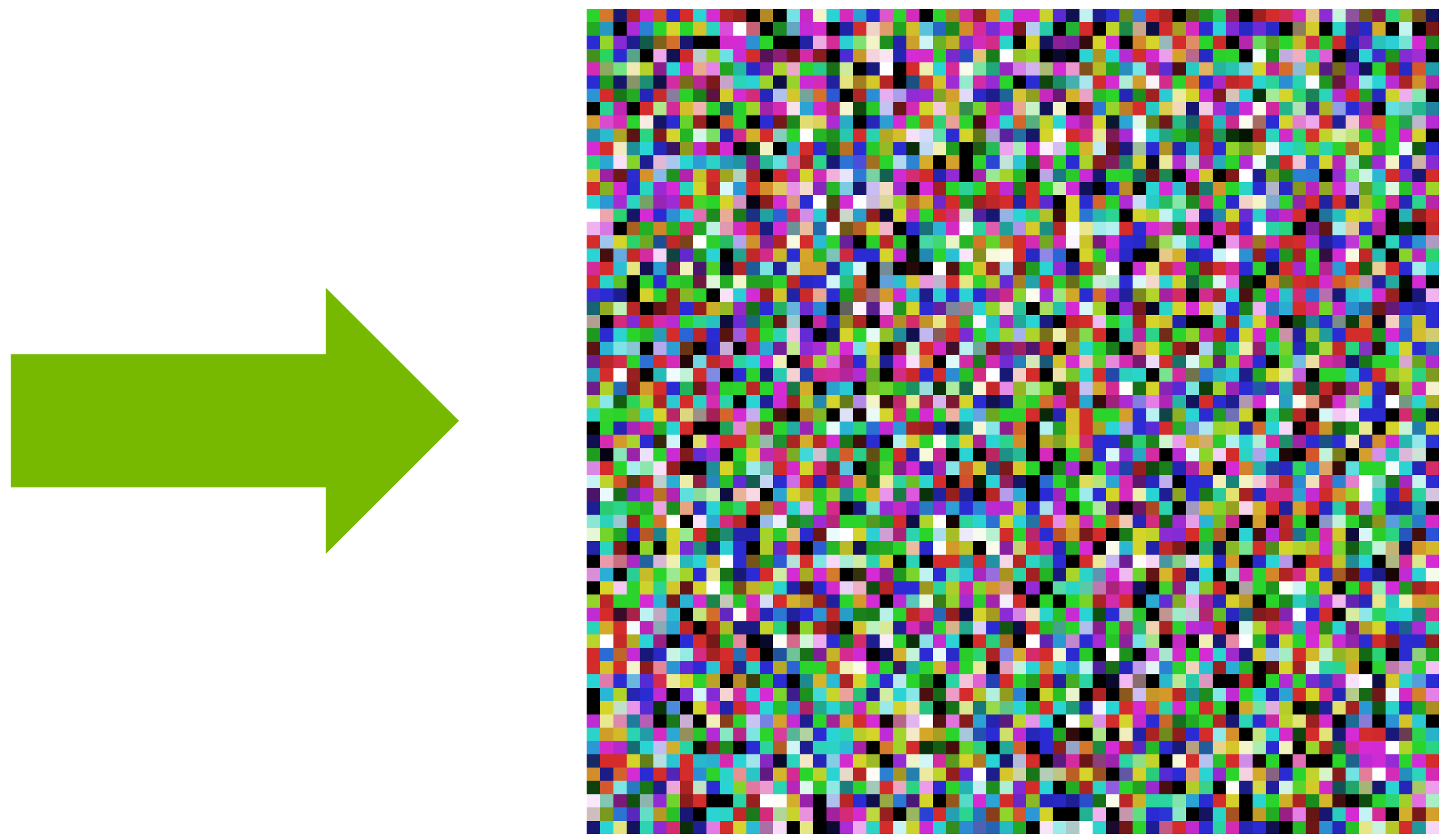
From Coin Flips to Bells

$$N(\mathrm{x}; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\mu = mean, a.k.a.\ average$ $\sigma = standard\ devaition, a.k.a.\ spread$

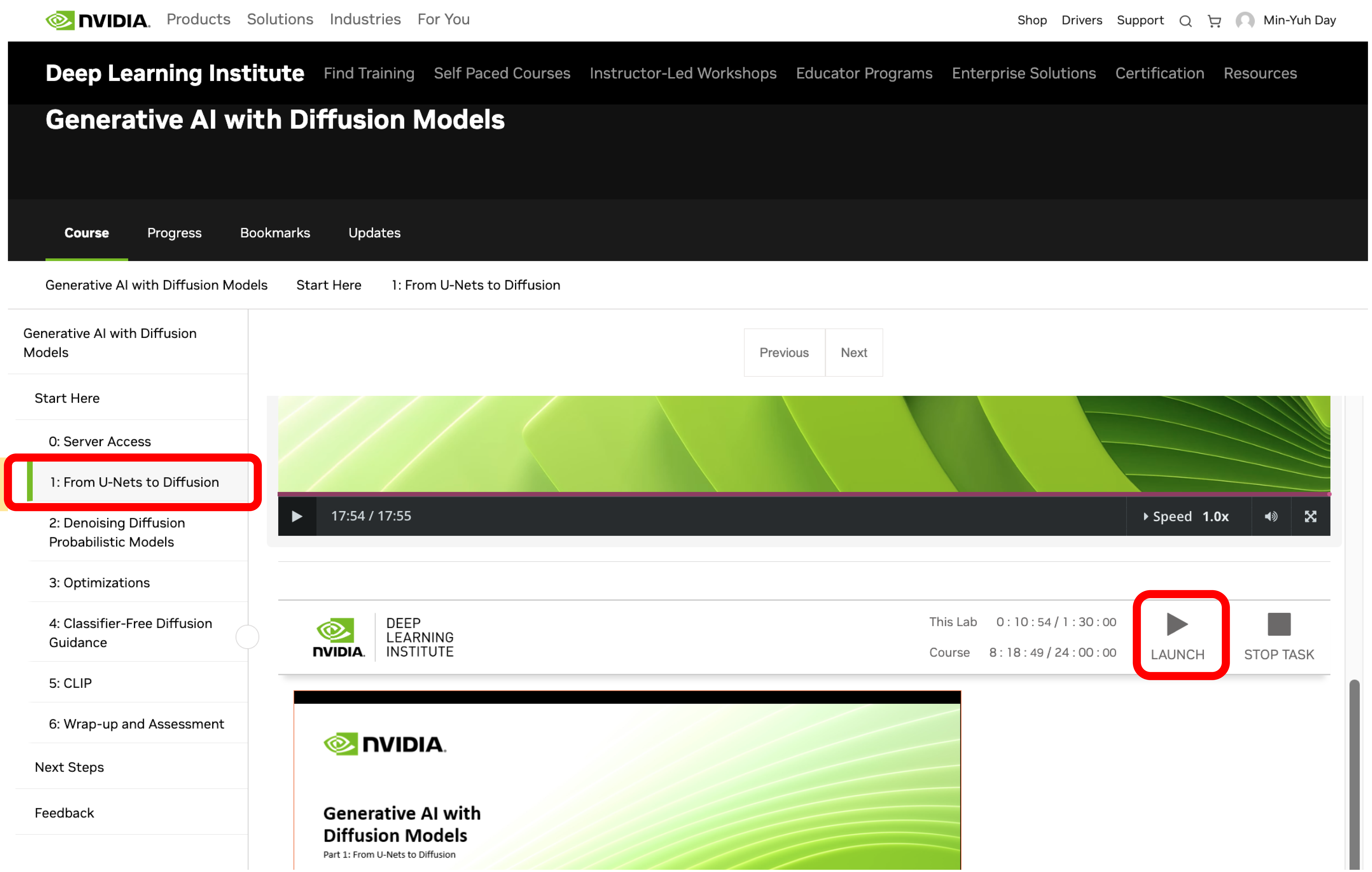
$$z = \frac{x\ - \mu}{\sigma}$$





Generative AI with Diffusion Models

1: From U-Nets to Diffusion

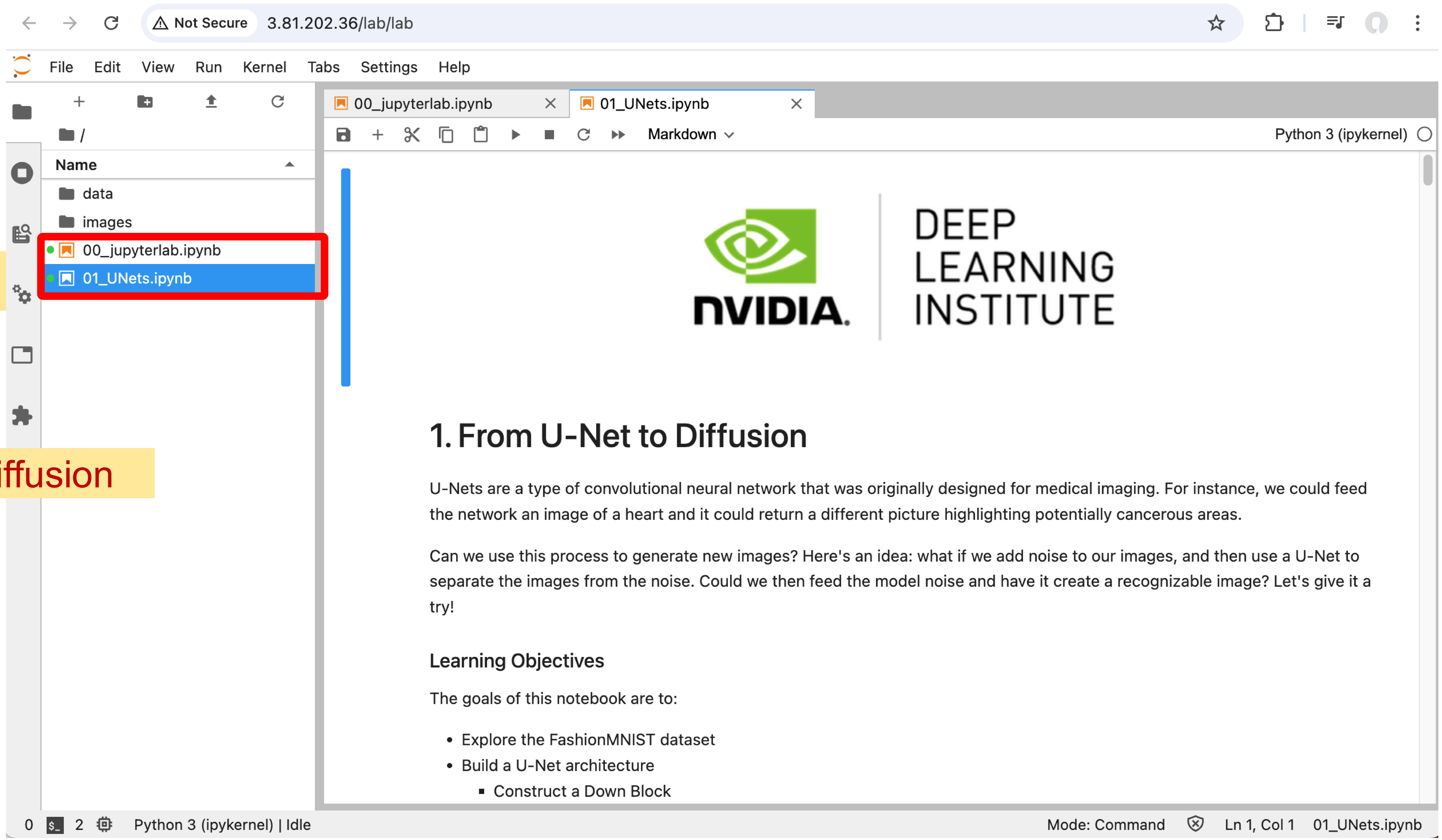


https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

01_UNets.ipynb

1: From U-Nets to Diffusion



https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

Part 2: Denoising Diffusion Probabilistic Models

Agenda

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- Part 3: Optimizations
- Part 4: Classifier-Free Diffusion Guidance
- Part 5: CLIP
- Part 6: Wrap-up & Assessment

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Diffusion Inspiration

Thermodynamic Origins

Deep Unsupervised Learning using Nonequilibrium Thermodynamics

Jascha Sohl-Dickstein JASCHA@STANFORD.EDU
Stanford University

Eric A. Weiss EAWEISS@BERKELEY.EDU
University of California, Berkeley

Niru Maheswaranathan NIRUM@STANFORD.EDU
Stanford University

Surya Ganguli SGANGULI@STANFORD.EDU
Stanford University

Abstract

A central problem in machine learning involves modeling complex data-sets using highly flexible families of probability distributions in which learning, sampling, inference, and evaluation

these models are unable to aptly describe structure in rich datasets. On the other hand, models that are *flexible* can be molded to fit structure in arbitrary data. For example, we can define models in terms of any (non-negative) function $\phi(\mathbf{x})$ yielding the flexible distribution $p(\mathbf{x}) = \frac{\phi(\mathbf{x})}{Z}$, where Z is a normalization constant. However, computing this

Green food coloring dissolving in a glass of water

Thermodynamic Origins

Denoising Diffusion Probabilistic Models

Getting Lost in the Noise

Denoising Diffusion Probabilistic Models

Jonathan Ho
UC Berkeley
jonathanho@berkeley.edu

Ajay Jain
UC Berkeley
ajayj@berkeley.edu

Pieter Abbeel
UC Berkeley
pabbeel@cs.berkeley.edu

Abstract

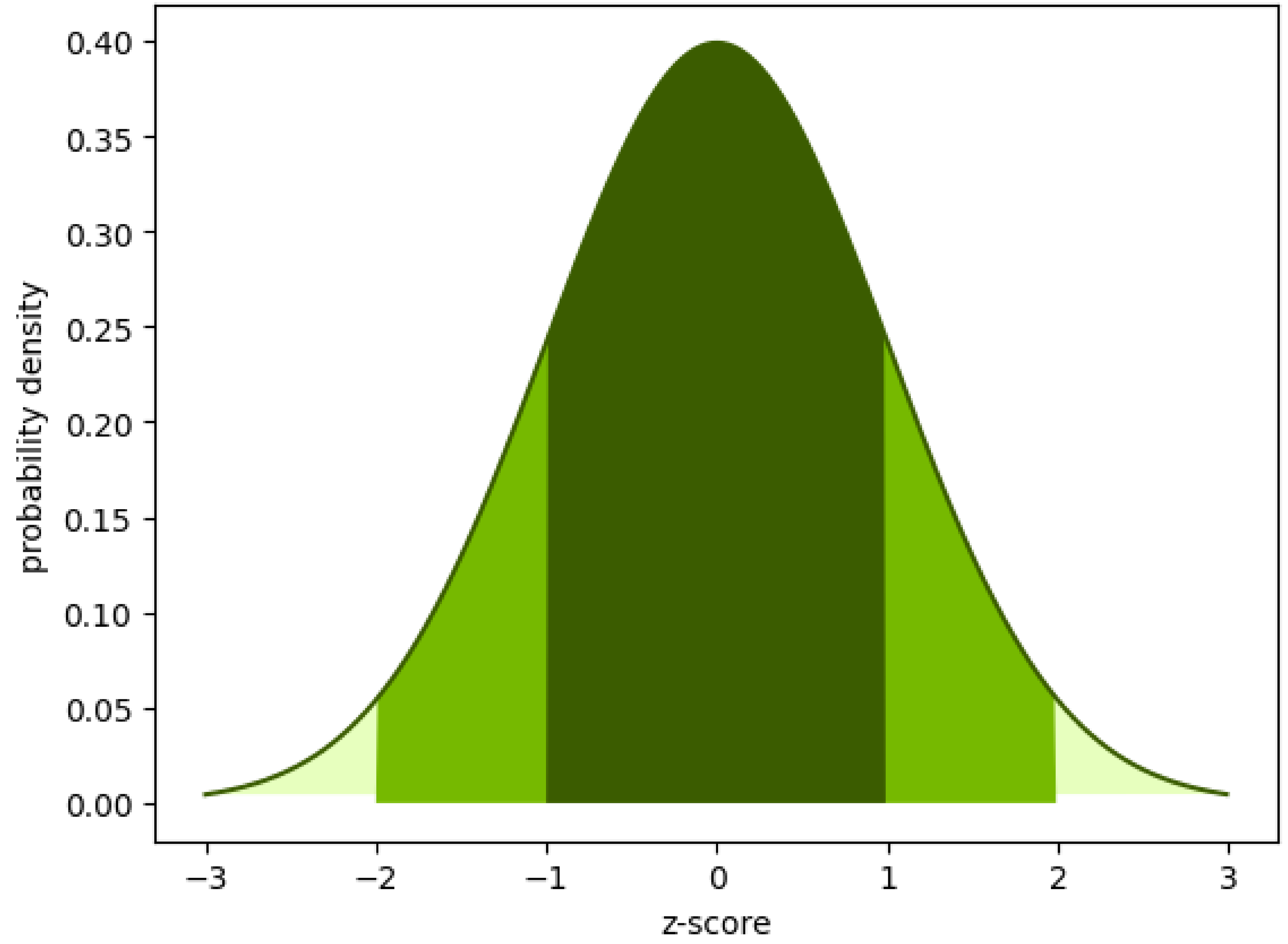
We present high quality image synthesis results using diffusion probabilistic models, a class of latent variable models inspired by considerations from nonequilibrium thermodynamics. Our best results are obtained by training on a weighted variational bound designed according to a novel connection between diffusion probabilistic models and denoising score matching with Langevin dynamics, and our models naturally admit a progressive lossy decompression scheme that can be interpreted as a

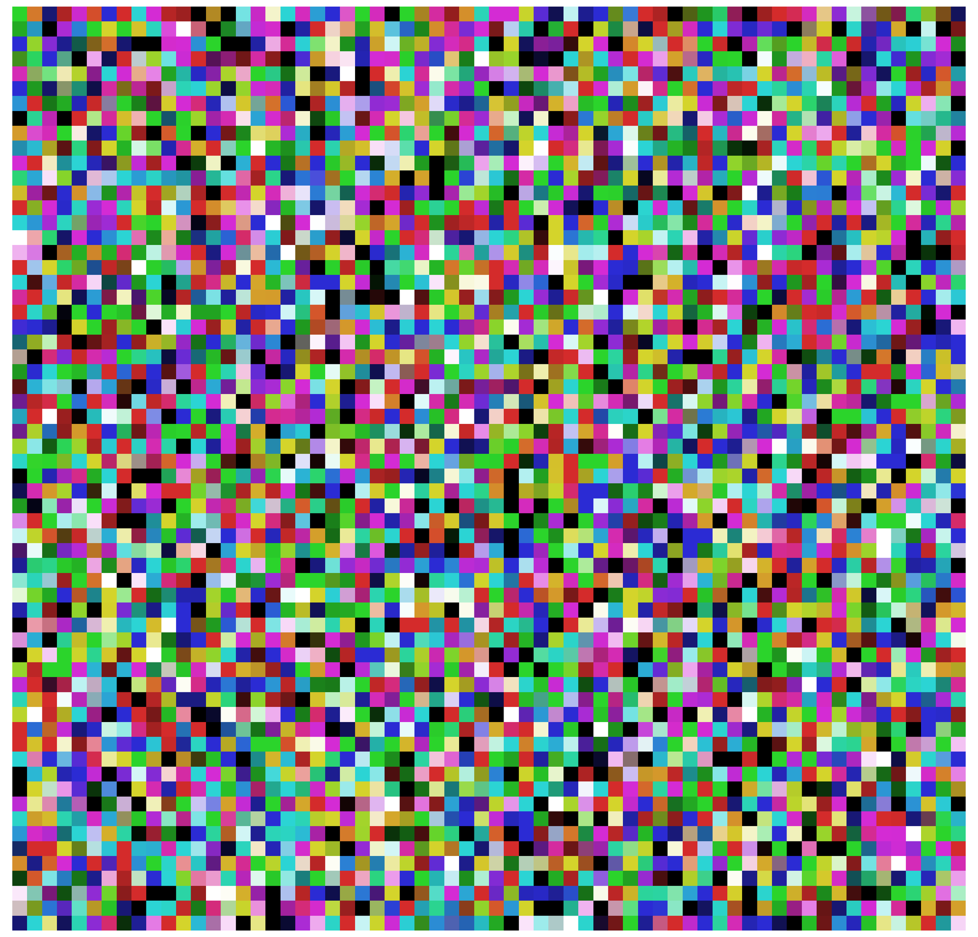
Forward Diffusion

Getting Lost in the Noise

Forward Diffusion

t = 0 | t = 1 | t = 2 | t = 3 | t = 4 | t = 5 | t = 6 | t = 7 | t = 8 | t = 9 | t = 10





$$N(\mathrm{x}; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Getting Lost in the Noise

Forward Diffusion

$$\beta = [0.0001, 0.0023, 0.0045, 0.0067, \dots, 0.0200]$$

t = 0 t = 1 t = 3 t = 4 t = 10

$T = 10$ — Total number of timesteps

$\alpha = 1 - \beta$ — Convenient shorthand

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Getting Lost in the Noise

Forward Diffusion


t = 0 t = 1 t = 2 t = 3 t = 4 t = 5 t = 6 t = 7 t = 8 t = 9 t = 10

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = N(\mathbf{x}_t; \sqrt{1-\beta_t} \cdot \mathbf{x}_{t-1}, \beta_t \cdot \mathbf{I})$$

Code:

```
noise = torch.randn_like(x_t)
x_t = torch.sqrt(1 - B[t]) * x_t
  + torch.sqrt(B[t]) * noise
```

Step 1:

- Generate noise from a standard normal distribution
- Multiply result by $\sqrt{\beta_t}$

Step 2:

- Multiply image at previous step by $\sqrt{1-\beta_t}$
- Add it to the result from step 1

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Reverse Diffusion

Getting Lost in the Noise

Reverse Diffusion

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

Difficult to calculate

Approximate average used to create $\mathbf{x}_t$

Getting Lost in the Noise

Reverse Diffusion

Getting Lost in the Noise

Reverse Diffusion

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = N(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\beta}_t \cdot \mathbf{I})$$

$$\tilde{\beta}_t = \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t$$

Use Bayes' Rule and …

Getting Lost in the Noise

Reverse Diffusion

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = N(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\beta}_t \cdot \mathbf{I})$$

$$\tilde{\mu}_t = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}}\epsilon_t\right)$$

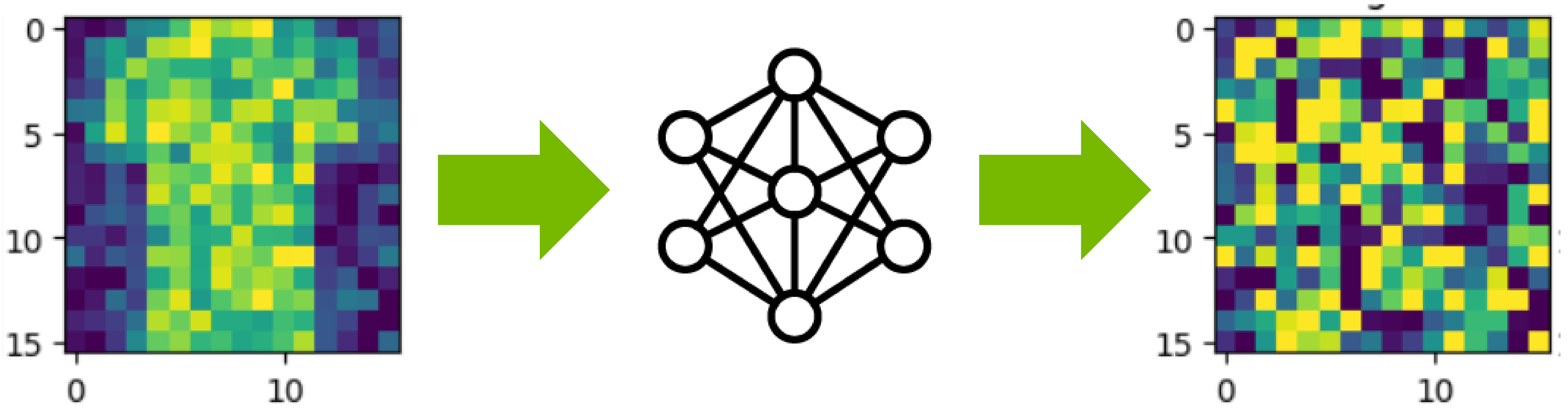
The noise added at time t
This is what the neural network will estimate

Model Training

Model Training

Training Data

Noisy image at time t



Noise added at time t

$$q(\mathbf{x}_t|\mathbf{x}_0) = N(\mathbf{x}_t; \sqrt{\bar{\alpha}_t} \cdot \mathbf{x}_0, (1 - \bar{\alpha}_t) \cdot \mathbf{I})$$

“Skip ahead” function

$t = 5$

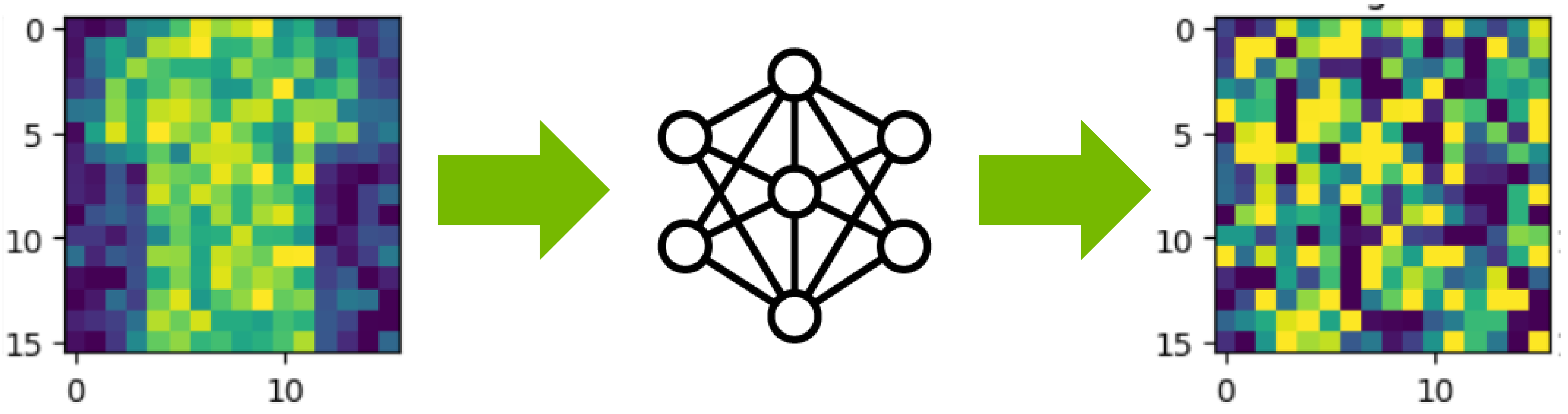
$t = 25$

$t = 100$

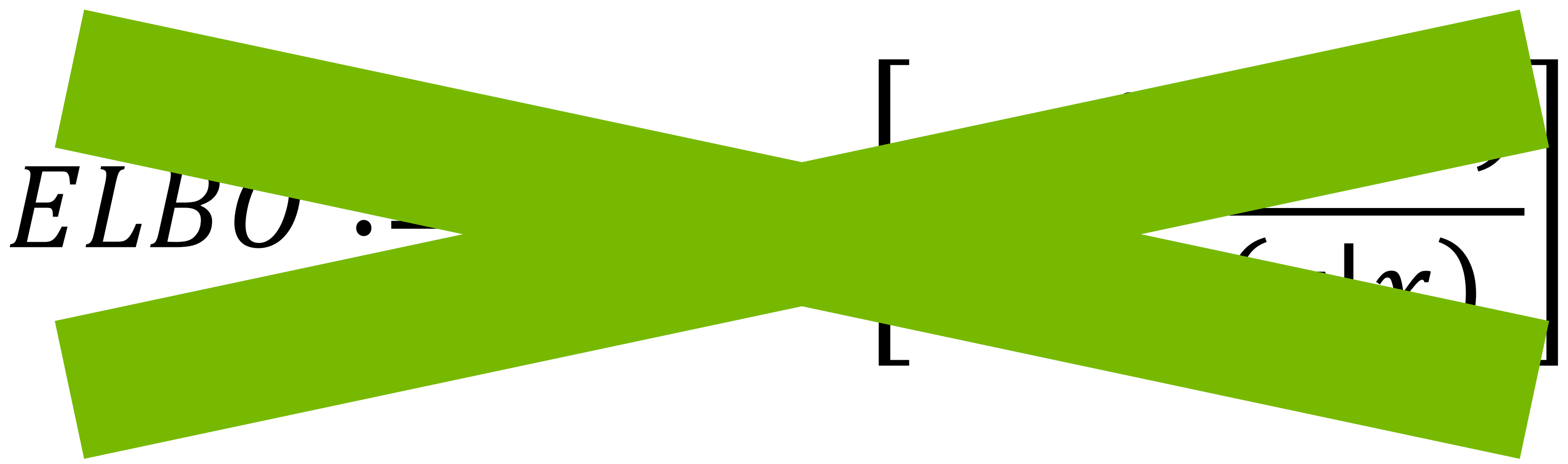
Model Training

ELBO Loss

Noisy image at time t



Noise added at time t



$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Model Architecture

The U-Net Architecture

Previous Lab

The U-Net Architecture

Next Lab

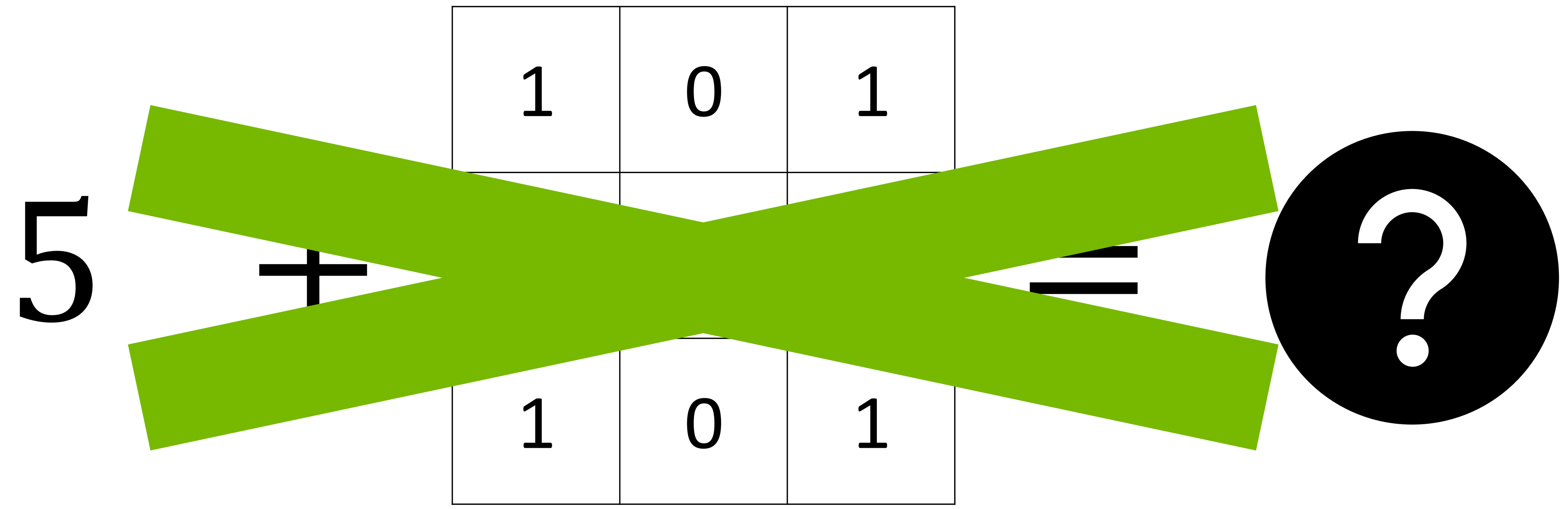
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

It's About Time

Adding Time

It’s About Time

Broadcasting



5

1	0	1
0	1	0
1	0	1

0 or 1

$3\ x\ 3$

Dimensions

It’s About Time

Broadcasting

$[[5]]$ +

1	0	1
0	1	0
1	0	1

$1 \; x \; 1$

Each dimension of 1 will repeat its element to align with destination

5	5	5
5	5	5
5	5	5

$3 \; x \; 3$

+

1	0	1
0	1	0
1	0	1

$3 \; x \; 3$

=

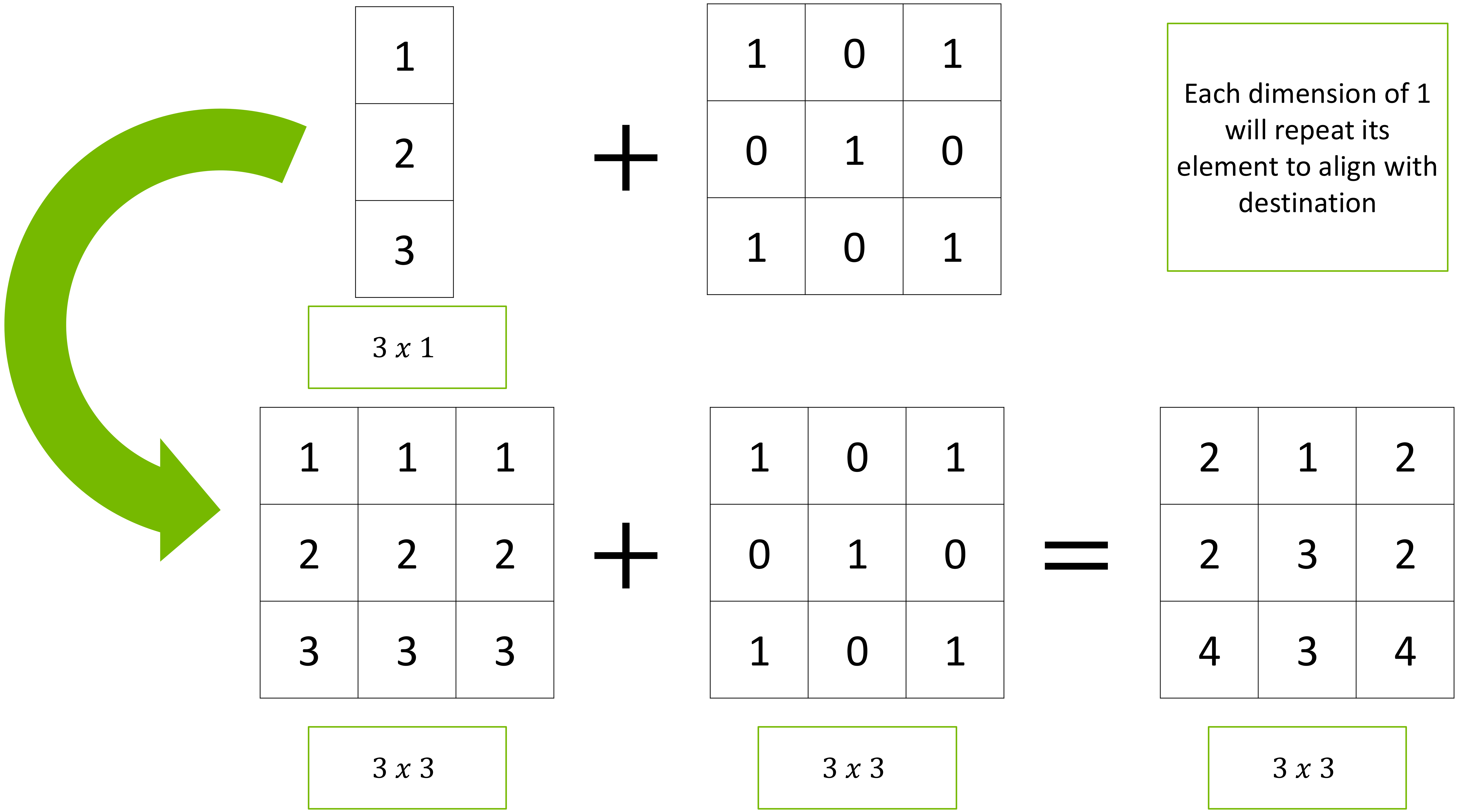
6	5	6
5	6	5
6	5	6

$3 \; x \; 3$

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

It's About Time

Broadcasting



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

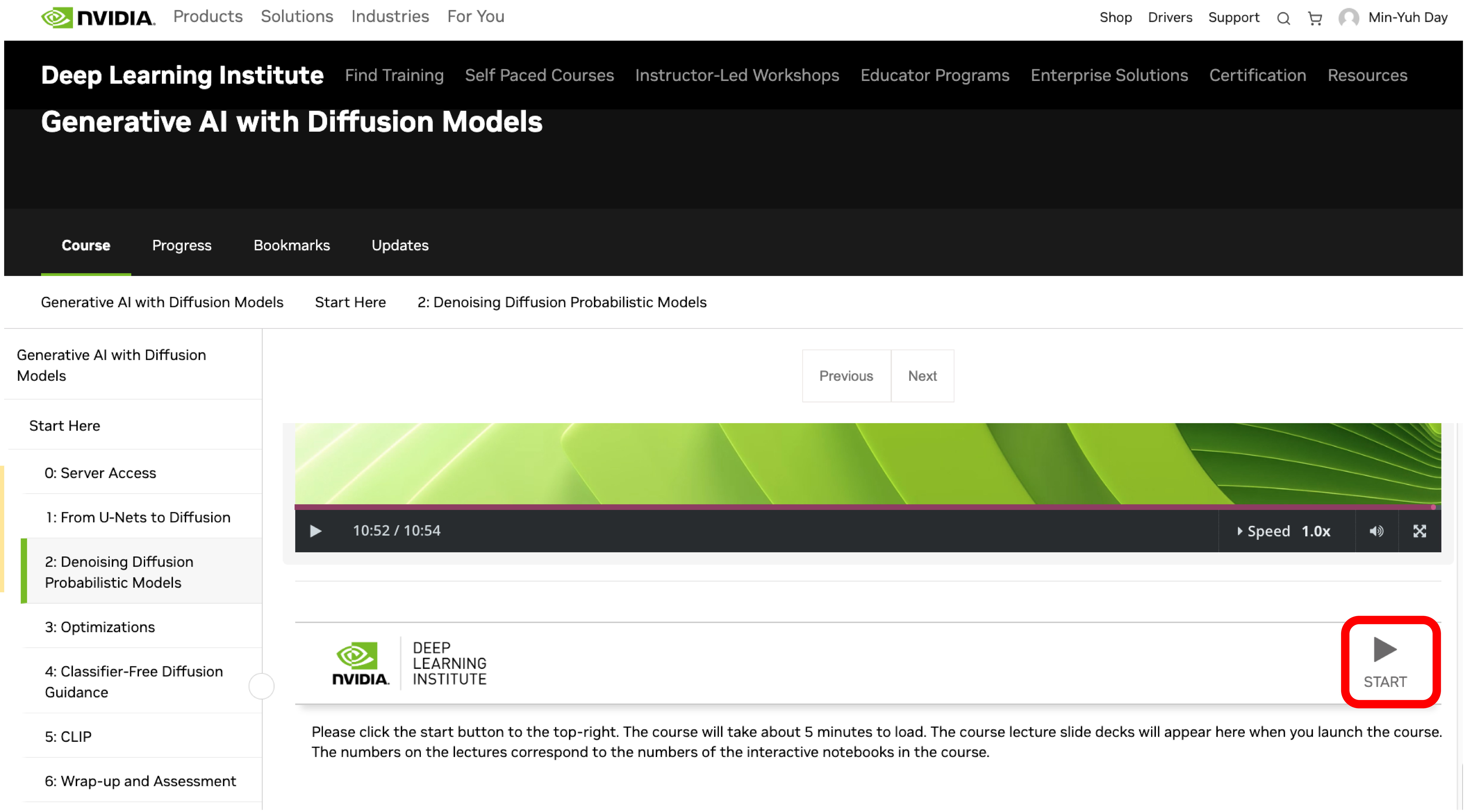
Let’s get started!

2: Denoising Diffusion Probabilistic Models

Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

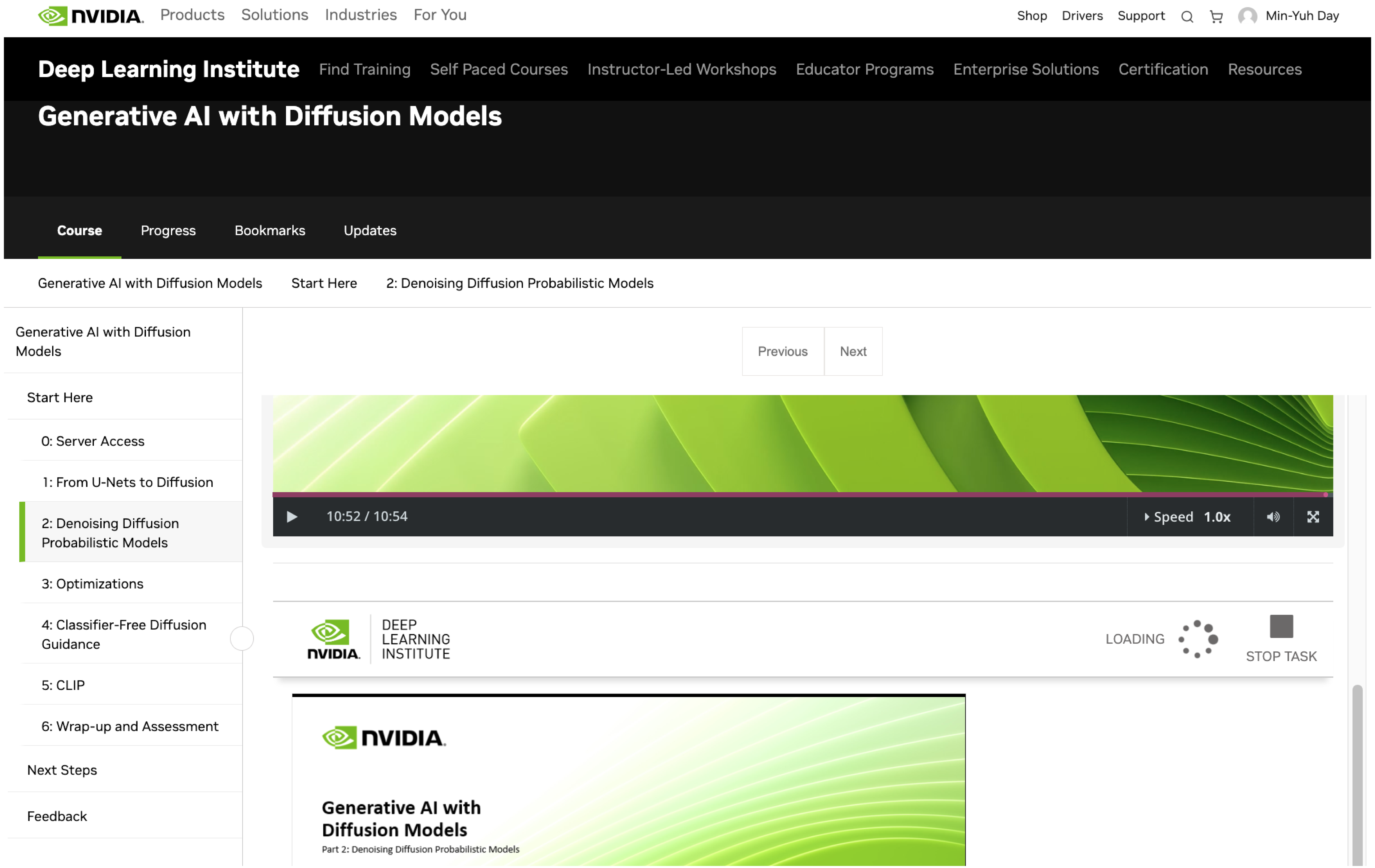
2: Denoising Diffusion Probabilistic Models



https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

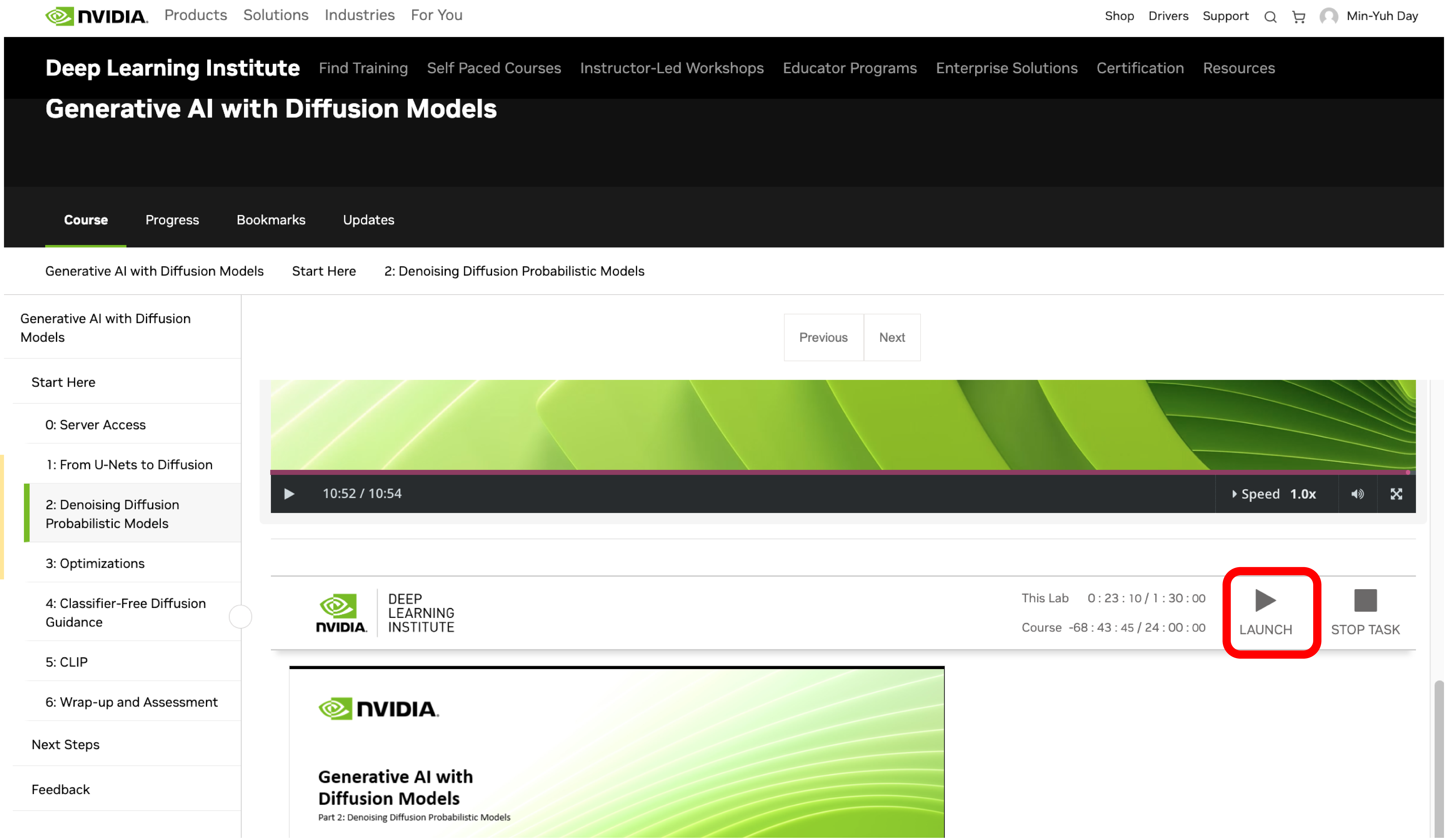
Generative AI with Diffusion Models

2: Denoising Diffusion Probabilistic Models



https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

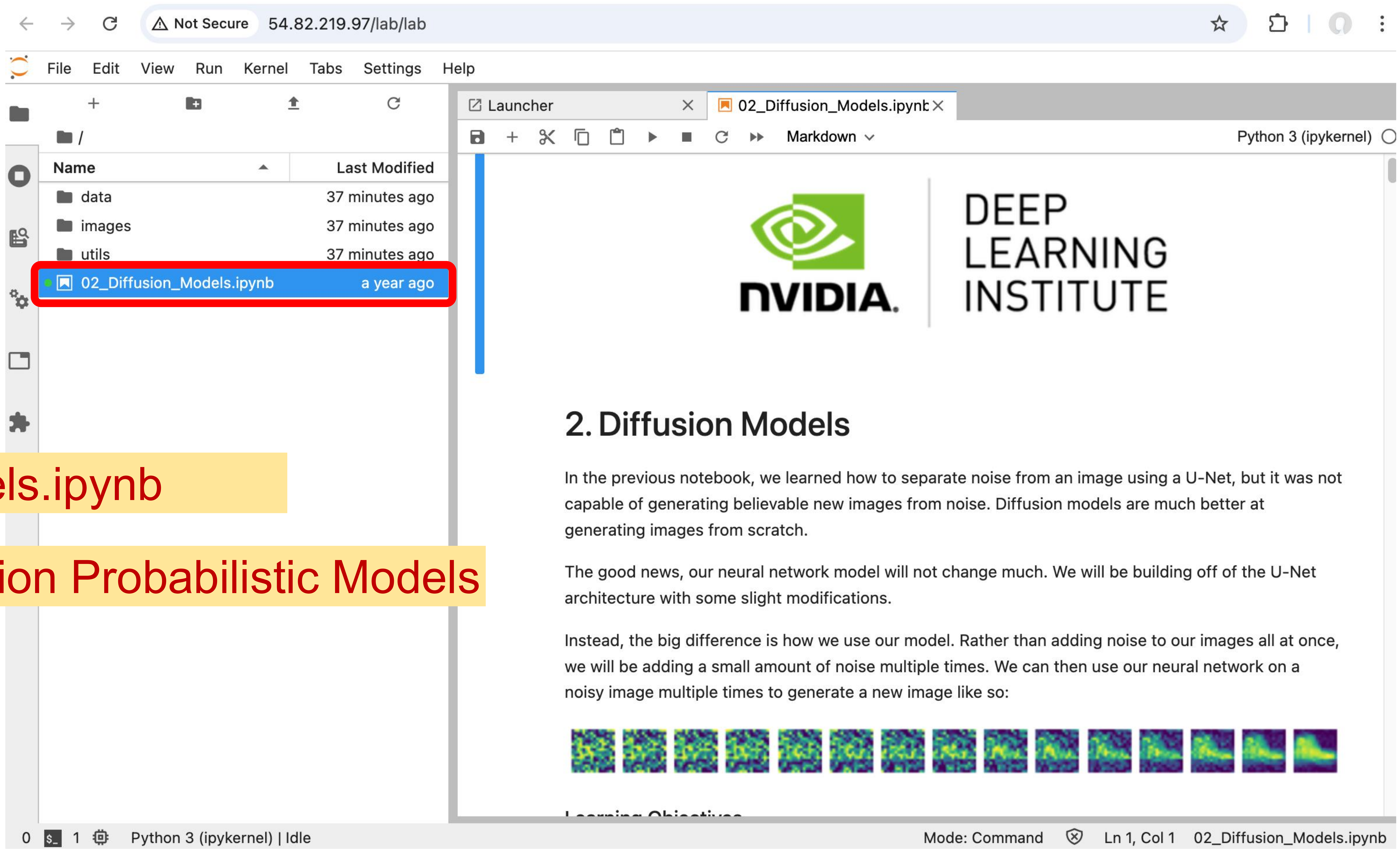
Generative AI with Diffusion Models



2: Denoising Diffusion Probabilistic Models

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

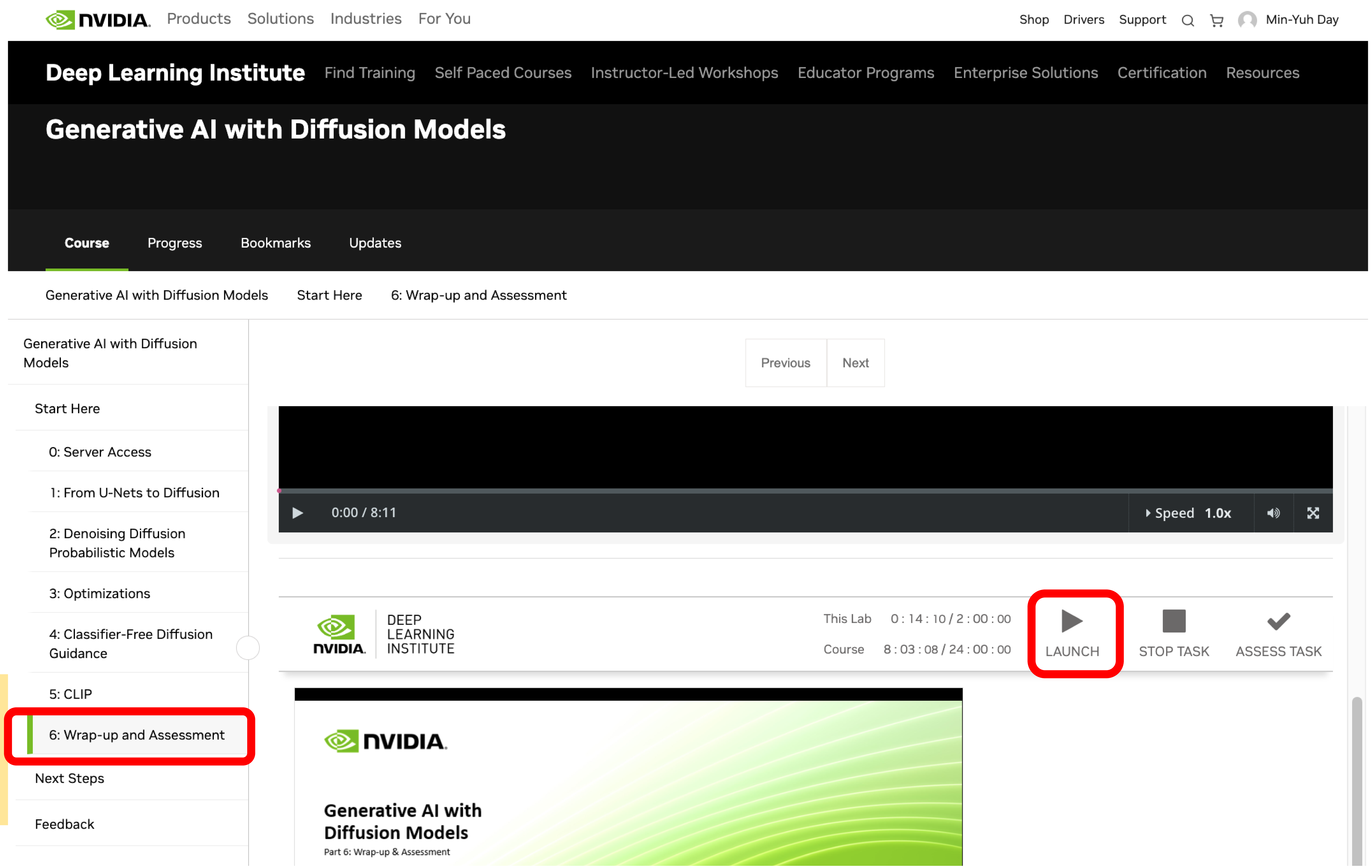


02_Diffusion_Models.ipynb

2: Denoising Diffusion Probabilistic Models

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

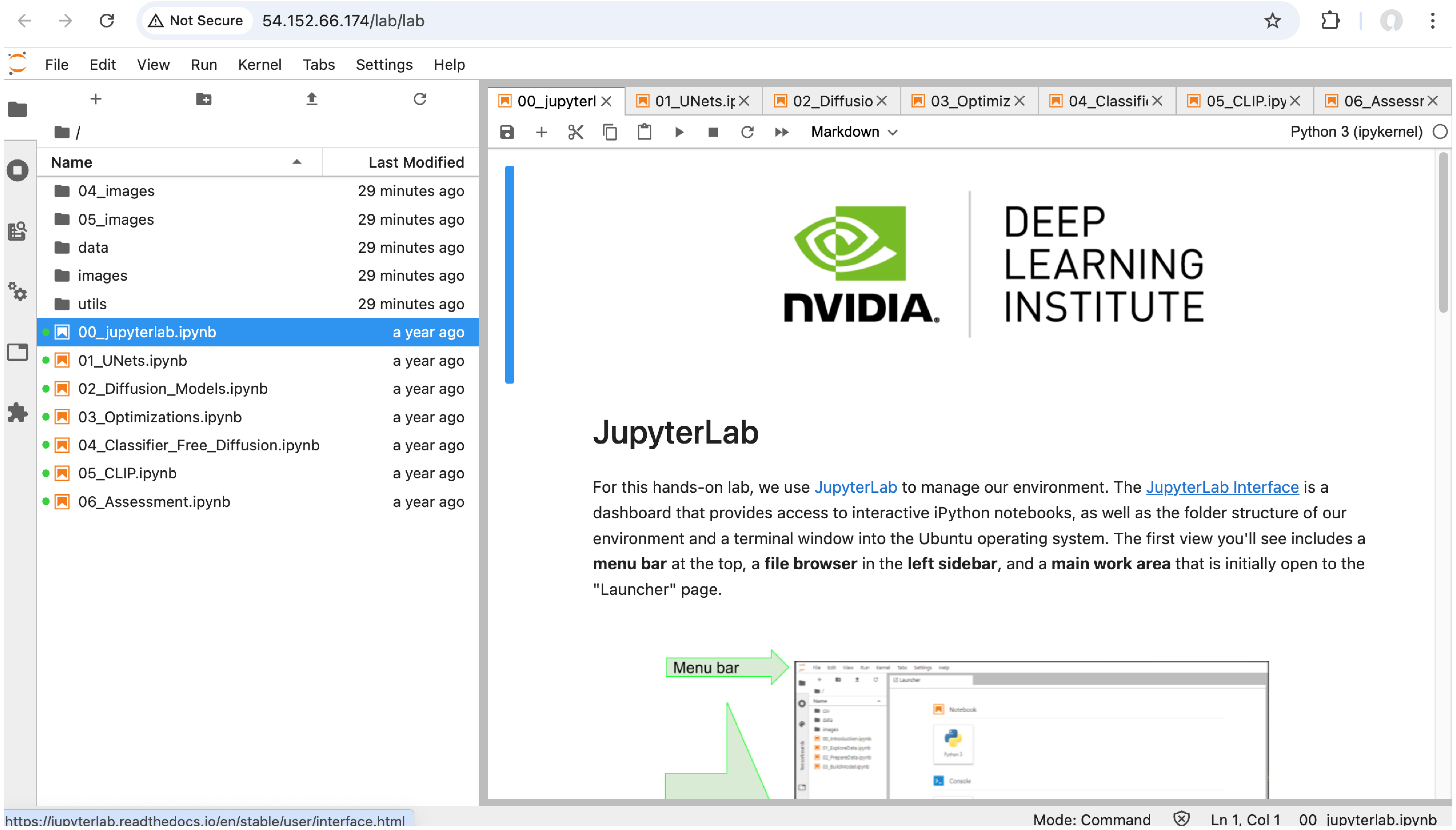
Generative AI with Diffusion Models



6: Wrap-up and Assessment

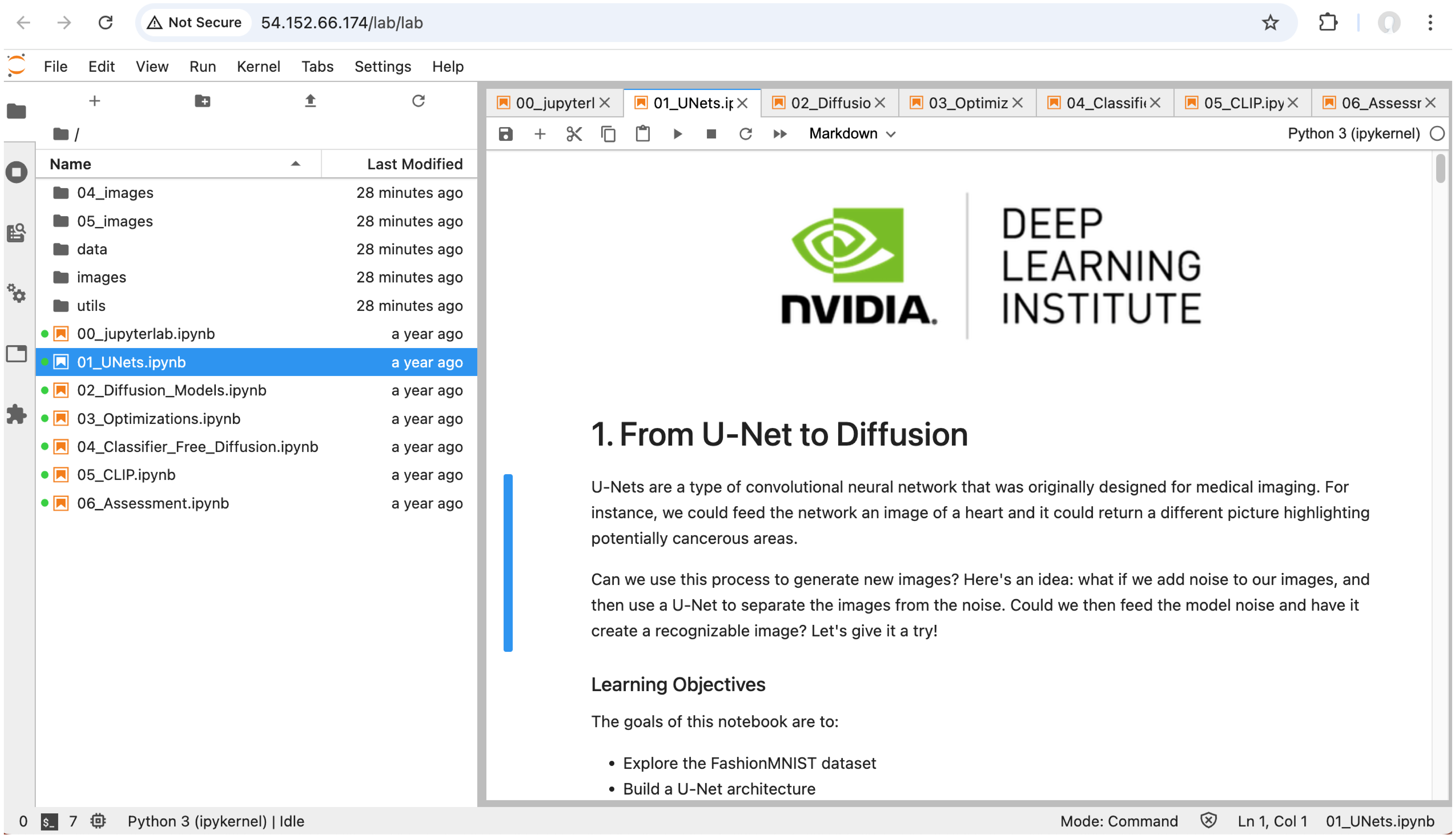
https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



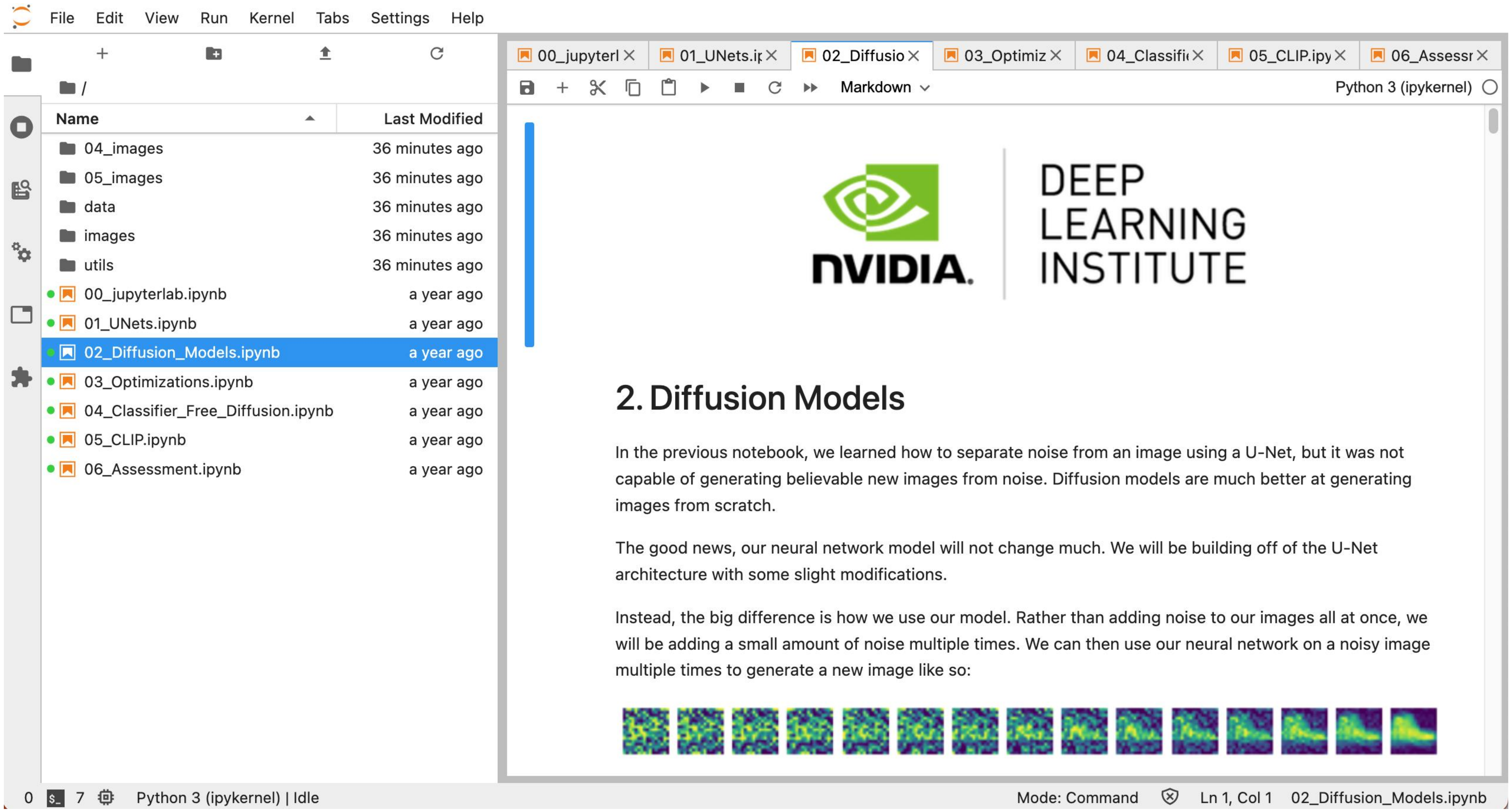
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



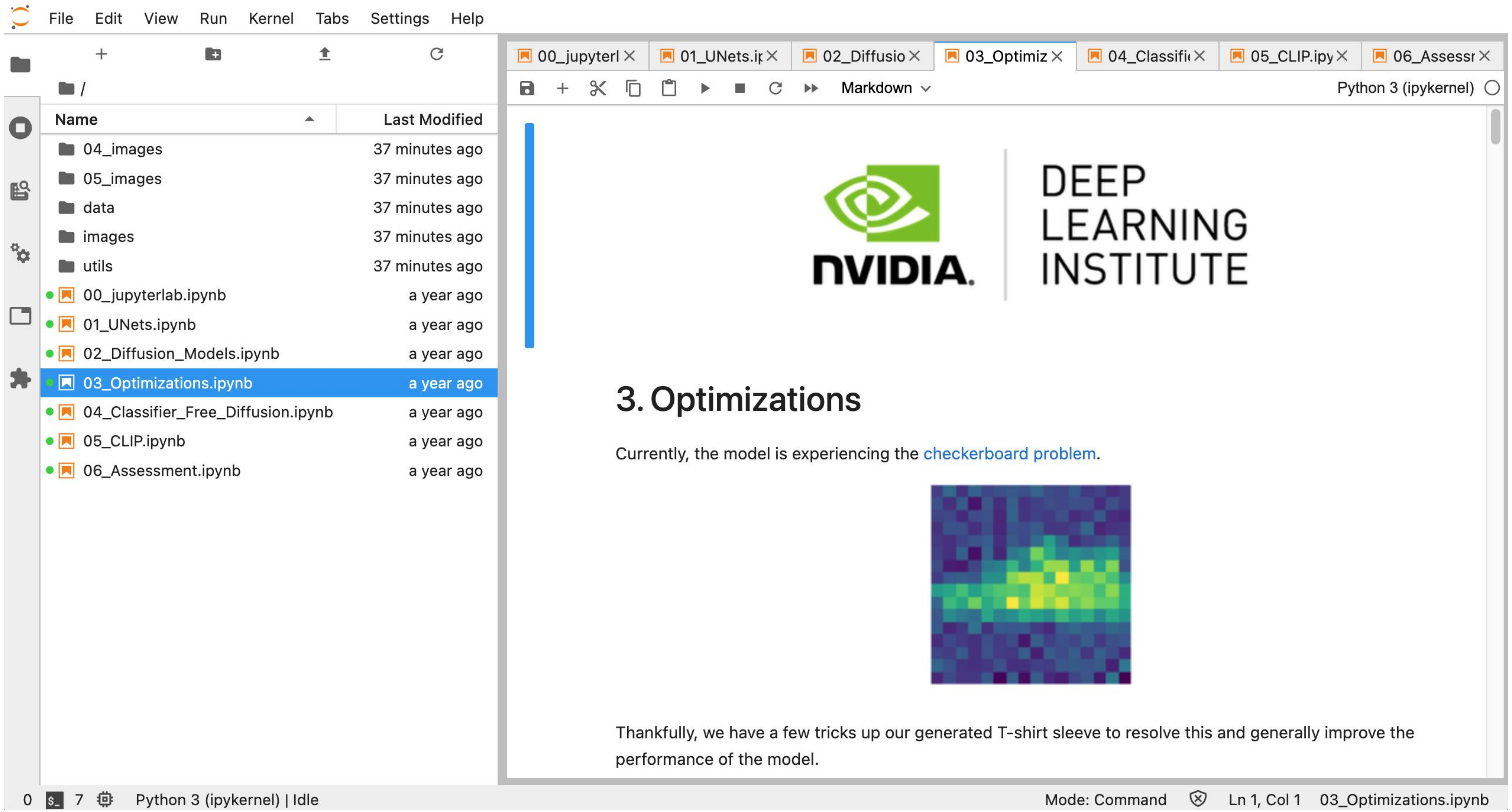
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



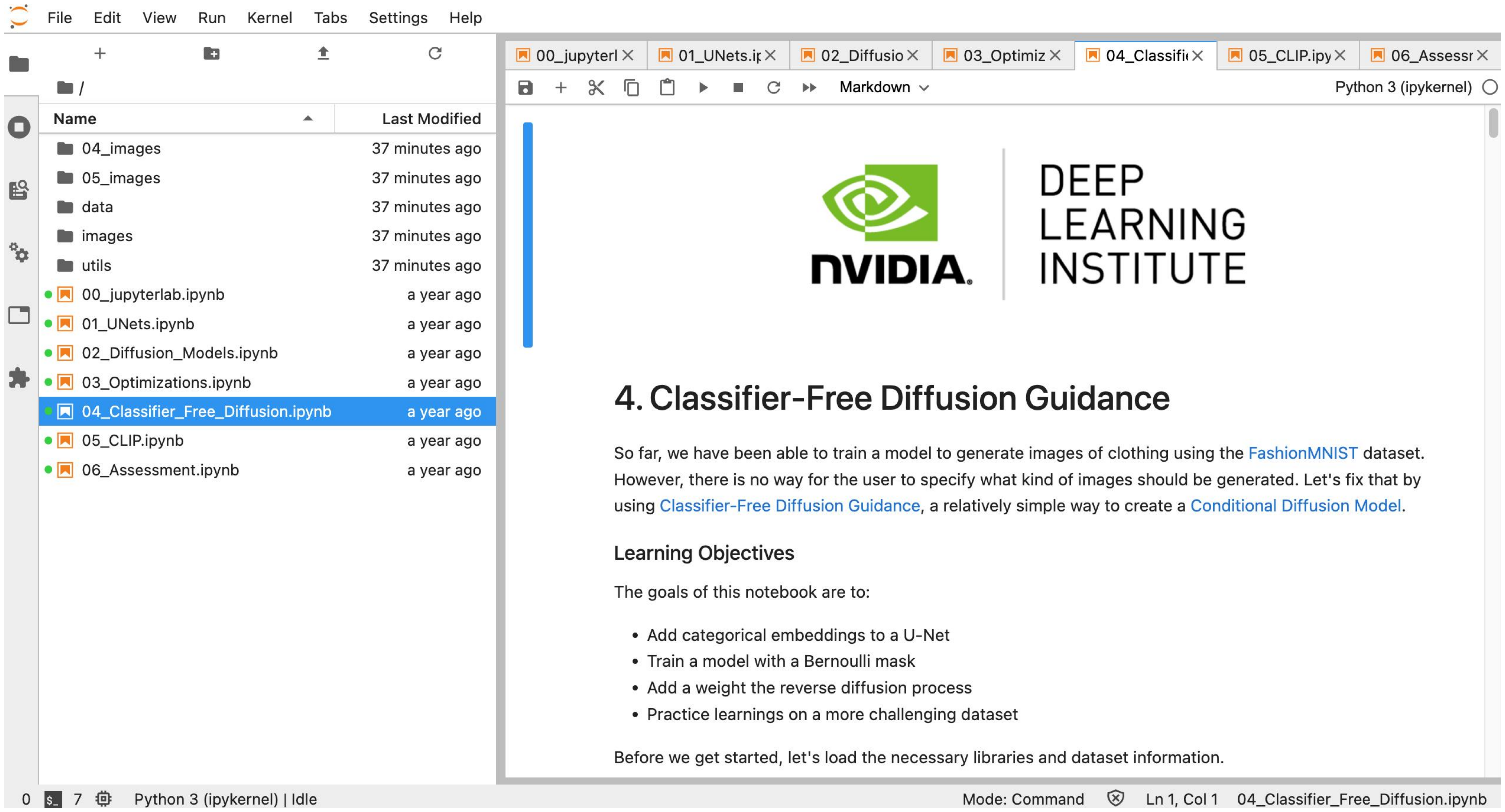
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



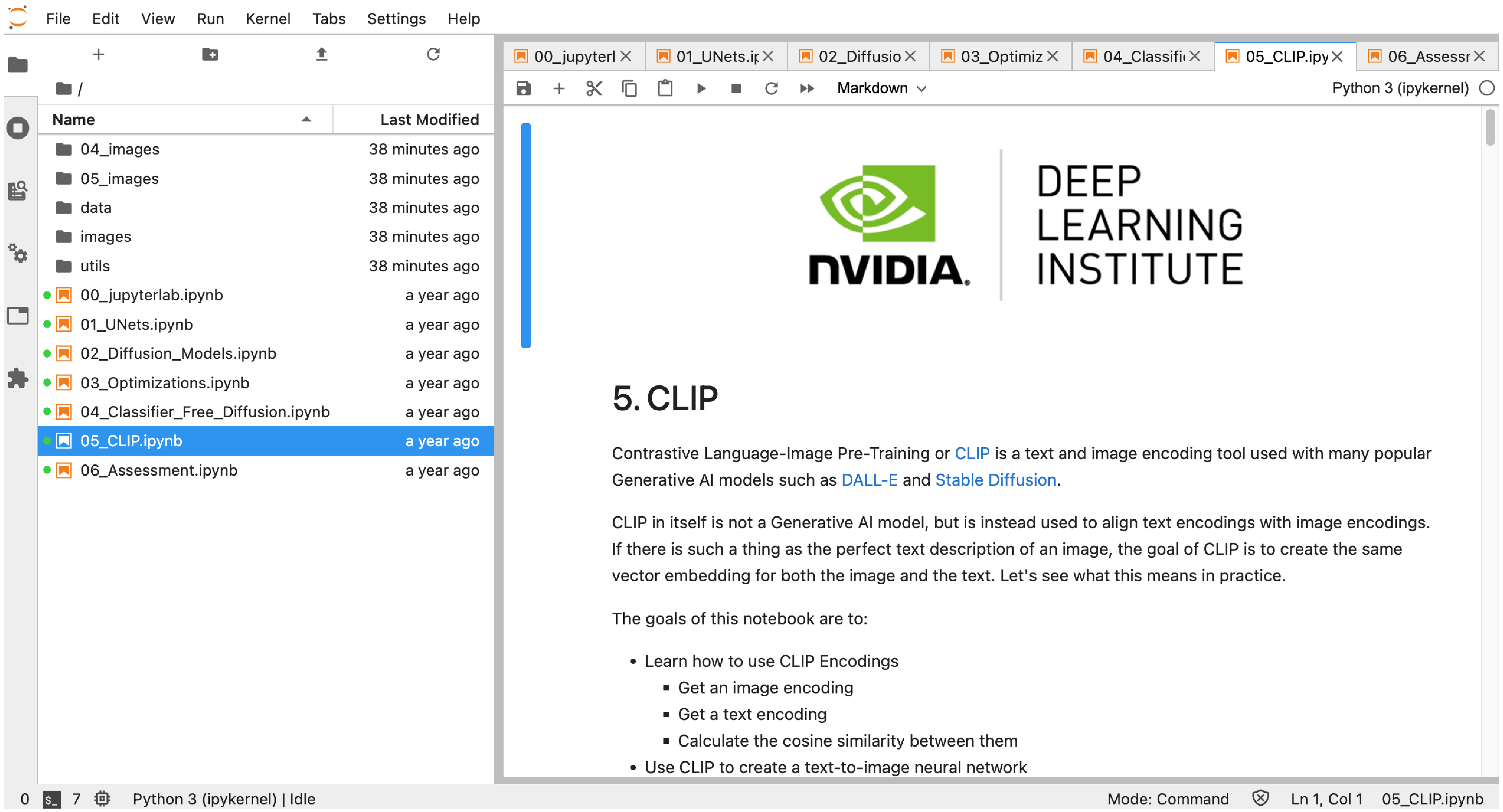
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



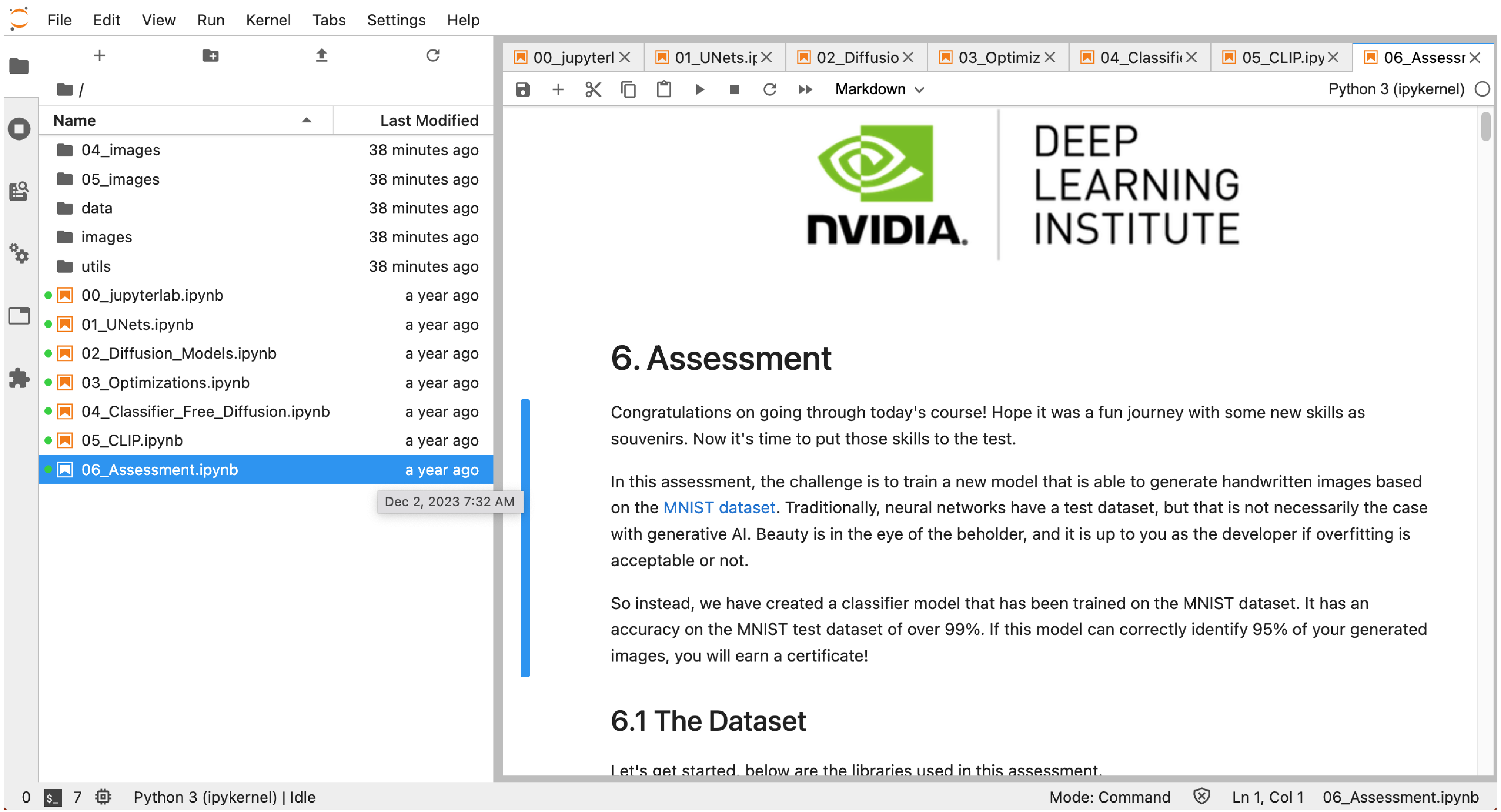
Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA Generative AI with Diffusion Models



Source: NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

NVIDIA

Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of

Generative AI with Diffusion Models

Greg Estes

Vice President, NVIDIA

Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

https://learn.nvidia.com/certificates?id=q3oo-oBhTQKtyCCote2E-Q

References

- NVIDIA DLI (2024), Building RAG Agents with LLMs, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1
- NVIDIA DLI (2024), Generative AI with Diffusion Models, https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1