Generative AI with Diffusion Models (NVIDIA)

Duration: 16 hours

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 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

- Build a U-Net to generate images from pure noise
- Improve the quality of generated images with the denoising diffusion process
- Control the image output with context embeddings
- Generate images from English text prompts using the Contrastive Language—Image Pretraining (CLIP) neural network

Topics Covered

- U-Nets
- Diffusion
- CLIP
- Text-to-image Models

Course Outline

From U-Net to Diffusion	• Build a U-Net architecture. • Train a model to remove noise from an image.
Diffusion Models	• Define the forward diffusion function. • Update the U-Net architecture to accommodate a timestep.

	• Define a reverse diffusion function.
Optimizations	• Implement Group Normalization. • Implement GELU. • Implement Rearrange Pooling. • Implement Sinusoidal Position Embeddings.
Classifier-Free Diffusion Guidance	• Add categorical embeddings to a U-Net. • Train a model with a Bernoulli mask.
CLIP	• Learn how to use CLIP Encodings. • Use CLIP to create a text-to-image neural network.