



ELECTIVE COURSE · 67030

# Building Vision and Audio AI with Deep Generative Models

Learn generative AI from the ground up · Build working AI systems

## AI Systems You Will Build

Vision and audio AI powered by foundation models, fine-tuned with data from state-of-the-art generative models.

## AI Models You Will Learn

VAEs · GANs · Stable Diffusion · ControlNet | Vision Transformers · CLIP · BLIP · CLAP | Prompt engineering for image and audio generation · Fine-tuning foundation models with PEFT.

## Industry Tools You Will Master

The same stack used by AI engineers at top companies: PyTorch, HuggingFace (Diffusers + Transformers), Ultralytics, OpenCV, Librosa, Torchaudio.

## Innovative AI Portfolio Project You Will Ship

Design, build, and present your own innovative generative AI system: a demo-ready project for your CV, job interviews, or grad school applications.

Hands-On Practice

Innovate by Doing

Cutting-Edge Models

Industry-Ready Toolbox