

Deep Generative Models

ESE 6800-004

Fall Semester 2024

René Vidal

Director of the Center for Innovation in Data Engineering and Science (IDEAS),
Rachleff University Professor, University of Pennsylvania
Amazon Scholar & Chief Scientist at NORCE



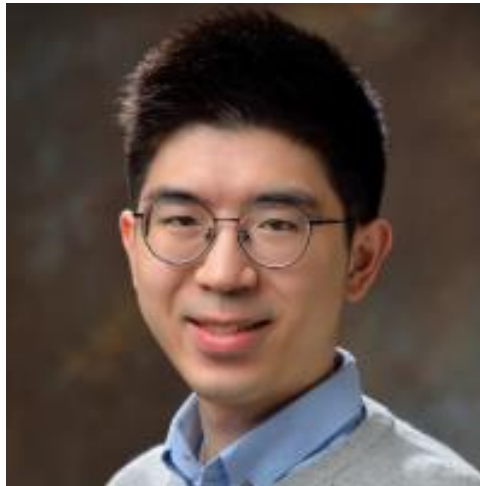
Course Information: Administrative



Instructor: René Vidal
vidalr@seas.upenn.edu

Lectures: TR 3:30pm-4:59pm in WLNT
401B (8/29 to 12/11)

Office Hours: Thursday 5pm-5:45pm,
WLNT 463C



TA: Ryan Chan
ryanckh@seas.upenn.edu



TA: Darshan Thaker
dbthaker@seas.upenn.edu



TA: Jinqi Luo
jinqiluo@seas.upenn.edu



TA: Tianjiao Ding
tjding@seas.upenn.edu



TA: Liangzu Peng
lpenn@seas.upenn.edu



TA: Kaleab A. Kinfu
kinfu@seas.upenn.edu

Office Hours

- Office Hours will start 09/02
- Professor René Vidal: Thursday 5pm-5:45pm, WLNT 463C
- TAs will rotate office hours every three weeks.
 - Darshan and Tianjiao: Tues 2-3pm. 09/03, 09/24, 10/15, 11/05
 - Ryan and Liangzu: Fri 12:30-1:30pm. 09/13, 10/04, 10/25, 11/15
 - Jinqi and Kaleab: Wed 3-4pm. 09/18, 10/09, 10/30, 11/20
- Location: WLNT 452C
- TAs will hold extra office hours on the week that homework is due.
- For extra OHs with René, feel free to email Sonia Castro Rodriguez soniacr@seas.upenn.edu

Course Website

- Please check our website for course updates:
 - <https://vidal-lab.github.io/dgm/>

University of Pennsylvania
Deep Generative Models
Fall 2024

HOME SCHEDULE LECTURES ASSIGNMENTS PROJECT MATERIALS

Deep Generative Models / Fall 2024

Updates

- New Lecture is up: Maximum Likelihood Estimation
- New Lecture is up: Introduction and Background

Course Description

Generative models have found widespread applications in science and engineering. Recent progress in deep learning has enabled the application of generative models to complex high-dimensional data such as images, videos, text and speech. This course will cover state-of-the-art deep generative models, including variational autoencoders (VAEs), auto-regressive models, diffusion models, and generative adversarial networks (GANs). The course will also illustrate various applications of deep generative models to image and video generation, text and speech generation, image captioning, text-to-image generation, and inverse problems.

Our course on [Canvas](#).

Previous Offerings

- [Spring 2024](#)

Instructor

René Vidal

Teaching Assistants

Ryan Chan

Darshan Thaker

Jinqi Luo

Tianjiao Ding

Liangzu Peng

Kaleab A. Kinfu

Schedule

Class Time: Tuesdays and Thursdays, 3:30 PM - 4:59 PM
Location: WLNT 401B

EVENT	DATE	DESCRIPTION	COURSE MATERIAL
Lecture	08/27/2024 Tuesday	Introduction and Background	Suggested Readings: <ul style="list-style-type: none">• History of AI• History of Generative Models• Advent of Deep Generative Models• Applications of Deep Generative Models• Discriminative vs. Generative Models
Lecture	08/29/2024 Thursday	Maximum Likelihood Estimation	Suggested Readings: <ul style="list-style-type: none">• Basics of Probability and Statistics• Basics of Information Theory• Basics of Maximum Likelihood Estimation• Basics of Optimization and Sampling
Lecture	09/03/2024 Tuesday	Latent Variable Models	Suggested Readings: <ul style="list-style-type: none">• Latent Variable Models
Lecture	09/05/2024 Thursday	Latent Variable Models	Suggested Readings: <ul style="list-style-type: none">• Probabilistic Principal Component Analysis (PPCA)• Application of PPCA to Generating Images of Faces with Variable Lighting
Lecture	09/10/2024 Tuesday	Latent Variable Models	Suggested Readings: <ul style="list-style-type: none">• Variational Auto-Encoders (VAEs)
Lecture	09/12/2024 Thursday	Latent Variable Models	Suggested Readings: <ul style="list-style-type: none">• Application of VAEs to Generating Images of Handwritten Digits and Faces

University of Pennsylvania
Philadelphia, PA 19104
United States

University of Pennsylvania
Philadelphia, PA 19104
United States

Course Information: Background

- Graduate-level course
- **Prerequisites:** Students should be comfortable with
 - **Multivariate calculus, linear algebra, and optimization:** gradient, Hessian, eigenvalue decomposition, singular value decomposition, first and second order conditions for minima/maxima, gradient descent, alternating minimization, Lagrange Multipliers
 - **Probability, statistics, information theory:** random variables, expectation, variance, covariance, maximum likelihood, expectation maximization, entropy, mutual information
 - **Programming:** Python
- Prior exposure to **machine learning** (e.g., ESE 4200, CIS 5190 or CIS 5200) is a plus.

Course Information: Books and Grading

- Text(s)/Required Materials:
 - Deep Generative Modeling, Jakub M. Tomczak, Springer Verlag, 2022 (<https://link-springer-com.proxy.library.upenn.edu/book/10.1007/978-3-030-93158-2>)
 - Generalized Principal Component Analysis, René Vidal, Yi Ma, Shankar Sastry, Springer Verlag, 2016 (<https://link-springer-com.proxy.library.upenn.edu/book/10.1007/978-0-387-87811-9>)
- Grading
 - Homework (60%): 3 mini-projects on different paradigms of generative models
 - Due roughly end of September, end of October, end of November
 - You will be given three (3) late days in total. You can freely arrange them across all homework.
 - Project (40%): text-to-image generation and editing
 - Due on exam day



Course Syllabus

- Introduction (Week 1)
 - History of AI, History of Generative Models (GMs), Advent of Deep GMs, Applications
- Background (Week 1 and 2)
 - Basics of Probability, Statistics, Information Theory, and Optimization
 - Discriminative vs Generative Models
 - Taxonomy of Generative Models
 - Maximum Likelihood Estimation
- Latent Variable Models (Week 3, 4 and 5)
 - Variational Inference and Expectation Maximization
 - Probabilistic Principal Component Analysis (PPCA)
 - Application of PPCA to Generating Images of Faces with Variable Lighting
 - Variational Auto-Encoders (VAEs)
 - Application of VAEs to Generating Images of Handwritten Digits

Course Syllabus

- Shallow Auto-regressive Models (Week 6)
 - Hidden Markov Models (HMMs)
 - Application of HMMs to Surgical Activity Recognition
 - Linear Dynamical Systems (LDSs)
 - Application of LDSs to Generating Videos of Dynamic Textures
- Deep Auto-regressive Models (Week 7, 8, and 9)
 - Recursive Neural Networks (RNNs)
 - Application to Speech Synthesis: WaveNet
 - Application to Image Captioning: RNN + VAE
 - Transformers (Transformer, Vision Transformer, Multimodal Transformers)
 - Application to Text Generation:
 - Application to Text-to-Image Generation: DALLE (auto-regressive model + VAE)

Course Syllabus

- Hybrid Models (Week 10)
 - Variational Autoencoders with Structured Encoders (GNNs, Transformers) and Priors (Autoregressive)
 - Applications to Image Captioning, Video Generation and Room Decoration.
- Diffusion Models (Week 11 and 12)
 - Denoising Diffusion Probabilistic Models (DDPM)
 - Diffusion Models with Latent Variables
 - Implementation details (Unet, StyleGAN)
 - Applications
 - Image Inpainting
 - Generating Images from Text: CLIP Guided Diffusion
- Generative Adversarial Networks (Week 13 and 14)
 - GANs, Wasserstein GAN, BiGAN, Style GAN
 - GAN Inversion
 - Application to Image Defenses, Image Denoising, Image Generation

Student Code of Conduct

- Read University Policy:
 - <https://grad.seas.upenn.edu/student-handbook/student-code-of-conduct/>
- You must not misrepresent someone else's work as your own. You can avoid this in two ways:
 - Do not use work (including code) from someone else.
 - Give proper credit if you do use someone else's work.
- Naturally, even if you give appropriate credit, you will only receive credit for your original work, so for this class you should stick with option #1.
- All cases of confirmed cheating/plagiarism will be reported to the Student Ethics Board.
- Homeworks and projects are strictly individual unless stated otherwise.

Exit Quiz

- To improve our class, we will have exit quiz that survey your opinions and comments for the course.
 - They are not mandatory and won't count towards your grade.
- Due every end of the week.

Registration & Auditing

- Number of registered students: ~50
- Number of students in the waiting list: ~40
- Capacity of WLNT 401B: ~ 54
- I'm really sorry since a larger room is not available, we will not be able to accept auditing students. All students must be registered.