Deep Generative Models (ESE 6450)

Fall Semester 2025

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

TA: Tianjiao Ding
tjding@seas.upenn.edu

TA: Liangzu Peng
lpenn@seas.upenn.edu

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

Lectures:
Tuesday & Thursday 3:30pm-4:59pm
at AGH 105 (8/26 to 12/4)

Office Hours:
Thursday 5pm-5:45pm, AGH 615

TA: Yuyan Ge
yyge@seas.upenn.edu

TA: Tadipatri Uday Kiran Reddy
ukreddy@seas.upenn.edu

TA: Kostas Emmanouilidis
ekostas@seas.upenn.edu

Office Hours

- Professor René Vidal: Thursday 5pm-5:45pm, AGH 615

- TAs will rotate office hours every three weeks.
  - Tianjiao and Uday: Wed 12-1pm. 09/03, 09/24, 10/15, 11/05
  - Liangzu and Kostas: Mon 2-3pm. 09/08, 10/29, 10/20, 11/10
  - Kaleab and Yuyan: Fri 9-10am. 09/19, 10/10, 10/31, 11/21

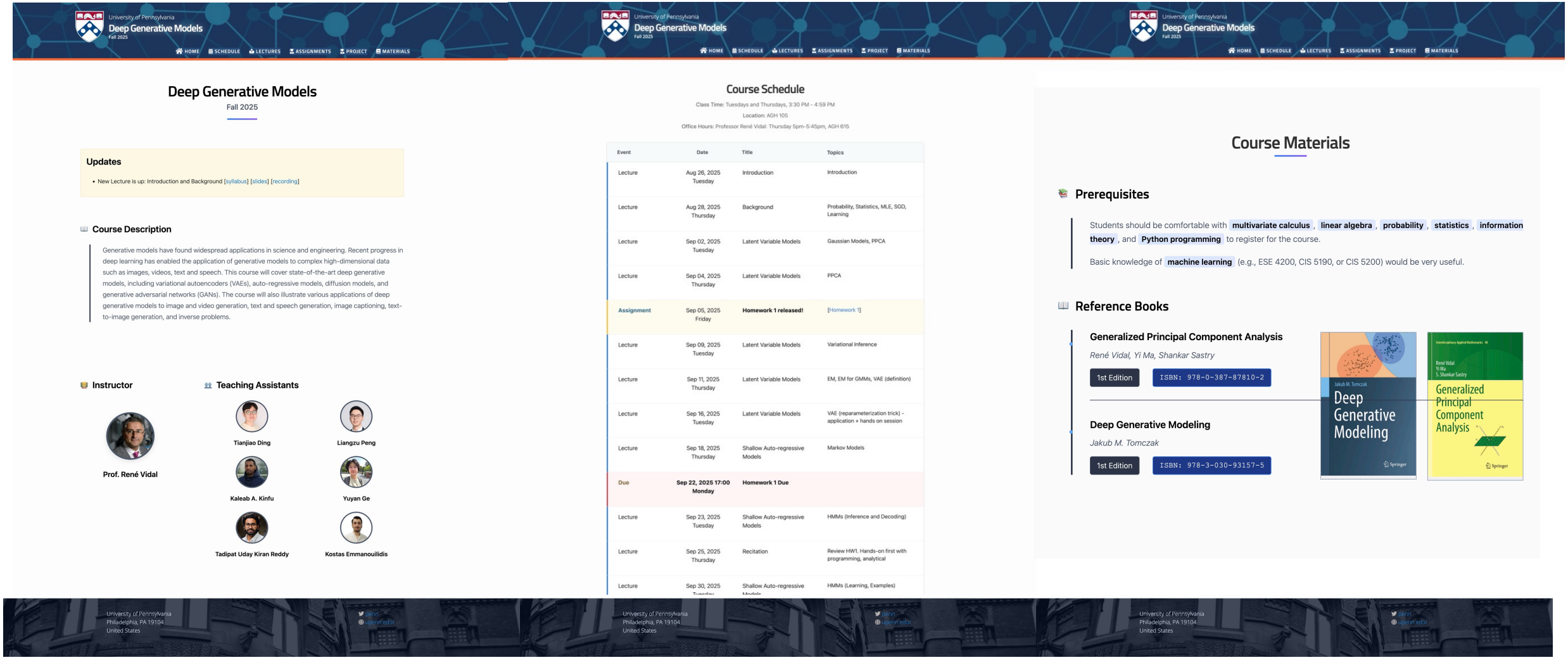
- Location: AGH 615

- 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/



Course Information: Background

- Graduate-level course

- **Prerequisites**: Students should be comfortable with
  - **Linear algebra:** eigenvalue decomposition, singular value decomposition
  - **Multivariate calculus, optimization**: gradient, Hessian, 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
    - Report due 12/2.
    - In-person poster presentation on 12/2 or 12/4.

Course Syllabus

- Introduction (Week 1)
  - History of AI, History of Generative Models (GMs), Advent of Deep GMs, Applications

- Background (Week 1, 2)
  - Basics of Linear Algebra, Probability, Statistics, Information Theory, and Optimization
  - Discriminative vs Generative Models
  - Taxonomy of Generative Models
  - Maximum Likelihood Estimation

- Latent Variable Models (Week 2, 3, 4)
  - 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 5, 6, 7, 8)
  - Markov Models & 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 8, 9, 10, 11, 12, 13)
  - Recursive Neural Networks (RNNs)
    - Application to Speech Synthesis (WaveNet) and Image Captioning (RNN + VAE)
  - Transformers (Transformer, Vision Transformer, Multimodal Transformers)
    - Application to Text Generation, and Text-to-Image Generation (DALLE: auto-regressive model + VAE)
  - Diffusion Models
    - Denoising Diffusion Probabilistic Models (DDPM)
    - Diffusion Models with Latent Variables
    - Implementation details (Unet, StyleGAN)
    - Applications to Image Inpainting, and Text-to-Image Generation (CLIP Guided Diffusion)

Course Syllabus

- Multimodal Models (Week 13, 14)
  - Vision-Language Models
  - ControlNet
  - Applications to Image Editing
- Project Presentations (Week 15)

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: ~62
- Number of students in the waiting list: ~85
- Capacity of AGH 105: ~72
- 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.