Generative AI

2025-2026

Penousal Machado

The Team

Theory

Penousal Machado

- machado@dei.uc.pt
- https://calendly.com/penousal/zoom-meeting
- https://videoconf-colibri.zoom.us/my/penousal

Practical Classes

João Correia
jncor@dei.uc.pt

Penousal Machado
machado@dei.uc.pt

Pedro Louro
pedrolouro@dei.uc.pt

Sancho Simões
sanchoamaralsimoes@gmail.com

Goals

Course Objectives

Knowledge:

- Comprehensive understanding of generative AI approaches
- Analysis of current challenges and opportunities in the field

Skills:

- Develop and adapt generative systems for real-world needs
- Apply generative AI across diverse domains through hands-on projects

Syllabus

Syllabus

1. **Introduction to Generative AI**
   - Overview and applications
   - Distinction between generative and discriminative models
   - Historical context and evolution
2. **Classical Generative AI Methods**
   - Rule-based and expert systems
   - Production systems
   - Constraint logic programming
   - Evolutionary
   - Markov chains (for generativity)
3. **Autoencoders**
   - Autoencoders the basics
   - Non-linear dimensionality reduction
   - Variational Autoencoders
4. **Generative Adversarial Networks**
5. **Diffusion Models**
6. **Generative ML Exploration Techniques**
   - Latent space exploration
   - Prompt generation and improvement
   - Optimisers
   - Quality-Diversity
7. **Text generation and large language models**
8. **Generative AI State of the Art**
   - Image generation and synthesis
   - Music and sound generation
   - Video generation and synthesis
   - Multimedia and hybrid models
9. **Ethical Considerations, Challenges, Opportunities.**

Course Evaluation

Assessment

- Group Projects 35%
  - 1st project 15%
  - 2nd project 20%
  - Group size ≤ 2
  - Oral defense
- Exam 65%
  - General questions 30-40%
  - Questions directly related with the projects: 20-30%

Bibliography

EVERYTHING EVERYWHERE ALL AT ONCE

Core Textbooks

Foster, D. (2023). Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play (2nd Ed.). O'Reilly Media.

Comprehensive: GANs, VAEs, Diffusion, Transformers with practical examples

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Foundational theory: Ch. 14 (Autoencoders), Ch. 20 (Deep Generative Models)

Prince, S.J.D. (2023). Understanding Deep Learning. MIT Press.

Modern treatment with diffusion models & transformers | **Free: udlbook.github.io**

Specialized Books

Langr, J. & Bok, V. (2019). GANs in Action: Deep Learning with Generative Adversarial Networks. Manning.

Practical GAN implementations with hands-on examples

Machado, P., Romero, J., & Greenfield, G. (Eds.) (2021). Artificial Intelligence and the Arts: Computational Creativity, Artistic Behavior, and Tools for Creatives. Springer.

Comprehensive coverage of AI techniques (evolutionary, neural, ML) in creative domains

Russell, S. & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th Ed.).

Classical AI: Logic, rule-based systems, production systems

Key Papers: Autoencoders & VAEs

Kingma, D.P. & Welling, M. (2014). Auto-Encoding Variational Bayes. ICLR.

Foundational VAE paper

Doersch, C. (2016). Tutorial on Variational Autoencoders. arXiv:1606.05908.

Excellent beginner-friendly tutorial

Key Papers: GANs

Goodfellow, I. et al. (2014). Generative Adversarial Networks. NeurIPS.

Original GAN paper

Radford, A. et al. (2016). Unsupervised Representation Learning with Deep Convolutional GANs. ICLR.

DCGAN architecture

Karras, T. et al. (2019). A Style-Based Generator Architecture for GANs (StyleGAN). CVPR.

State-of-the-art image synthesis

Key Papers: Diffusion Models

Ho, J. et al. (2020). Denoising Diffusion Probabilistic Models. NeurIPS.

DDPM - foundational diffusion paper

Dhariwal, P. & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. NeurIPS.

Showed diffusion superiority over GANs

Rombach, R. et al. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. CVPR.

Stable Diffusion architecture

Key Papers: Quality-Diversity

Pugh, J.K. et al. (2016). Quality Diversity: A New Frontier for Evolutionary Computation. Frontiers in Robotics and AI.

Overview of QD algorithms and applications

Mouret, J.-B. & Clune, J. (2015). Illuminating the Search Space by Mapping Elites. arXiv:1504.04909.

MAP-Elites algorithm

Cully, A. & Demiris, Y. (2017). Quality and Diversity Optimization: A Unifying Modular Framework. IEEE TEVC.

Unified QD framework

Key Papers: Text & LLMs

Vaswani, A. et al. (2017). Attention is All You Need. NeurIPS.

Transformer architecture

Brown, T. et al. (2020). Language Models are Few-Shot Learners (GPT-3). NeurIPS.

Demonstrated few-shot learning capabilities

OpenAI (2023). GPT-4 Technical Report. arXiv:2303.08774.

Multimodal capabilities

Key Papers: Image Generation

Ramesh, A. et al. (2022). Hierarchical Text-Conditional Image Generation with CLIP Latents (DALL-E 2). arXiv:2204.06125.

Text-to-image with CLIP guidance

Saharia, M. et al. (2022). Photorealistic Text-to-Image Diffusion Models (Imagen). NeurIPS.

High-fidelity text-to-image generation

Rombach, R. et al. (2022). Stable Diffusion. CVPR.

Open-source latent diffusion

Key Papers: Audio & Video

Dhariwal, P. et al. (2020). Jukebox: A Generative Model for Music. arXiv:2005.00341.

Huang, Q. et al. (2023). Make-An-Audio: Text-To-Audio with Prompt-Enhanced Diffusion. ICML.

Ho, J. et al. (2022). Imagen Video: High Definition Video Generation with Diffusion Models. arXiv:2210.02303.

Singer, U. et al. (2023). Make-A-Video: Text-to-Video without Text-Video Data. ICLR.

Ethics & Challenges

Mitchell, M. (2019). Artificial Intelligence: A Guide for Thinking Humans. Farrar, Straus and Giroux.

Accessible overview of AI capabilities, limitations, and societal implications

Vallor, S. (2024). The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking. Oxford University Press.

Philosophical perspective on AI's impact on human identity and agency

Bender, E.M. et al. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT.

Environmental and social concerns of large models

Weidinger, L. et al. (2021). Ethical and social risks of harm from Language Models. arXiv:2112.04359.

Comprehensive taxonomy of LM risks

Bommasani, R. et al. (2021). On the Opportunities and Risks of Foundation Models. arXiv:2108.07258.

Holistic analysis of foundation models

Online Resources & Tutorials

Hugging Face Course
huggingface.co/course — Practical transformers and LLMs

Lil'Log (Lilian Weng's blog)
lilianweng.github.io — Deep dives: VAEs, GANs, Diffusion, Transformers

Distill.pub
distill.pub — Visual explanations of ML concepts

Papers with Code
paperswithcode.com — papers with implementations

The Illustrated Transformer (Jay Alammar)
jalammar.github.io/illustrated-transformer


Stability Emerges