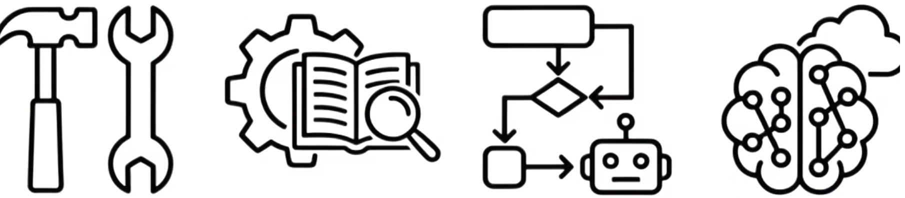
Deep Learning
Advanced Models & Methods

Prof. Antonino Furnari (antonino.furnari@unict.it)
Corso di Laurea Magistrale in Informatica
Dip. di Matematica e Informatica
Università di Catania

Introduction to the Course

LM-18 COURSE

Welcome to Deep Learning - Advanced Models and Methods



The Paradigm Shift

Previous courses: learned the **standard architectures**

This course: learn to **build the engine**

Your Transformation

From: Practitioner

To: Deep Learning Engineer

Building a rich toolbox of techniques

Who Am I?

Prof. Antonino Furnari

Contacts:

- http://antoninofurnari.github.io
- antonino.furnari@unict.it

Office Hours:

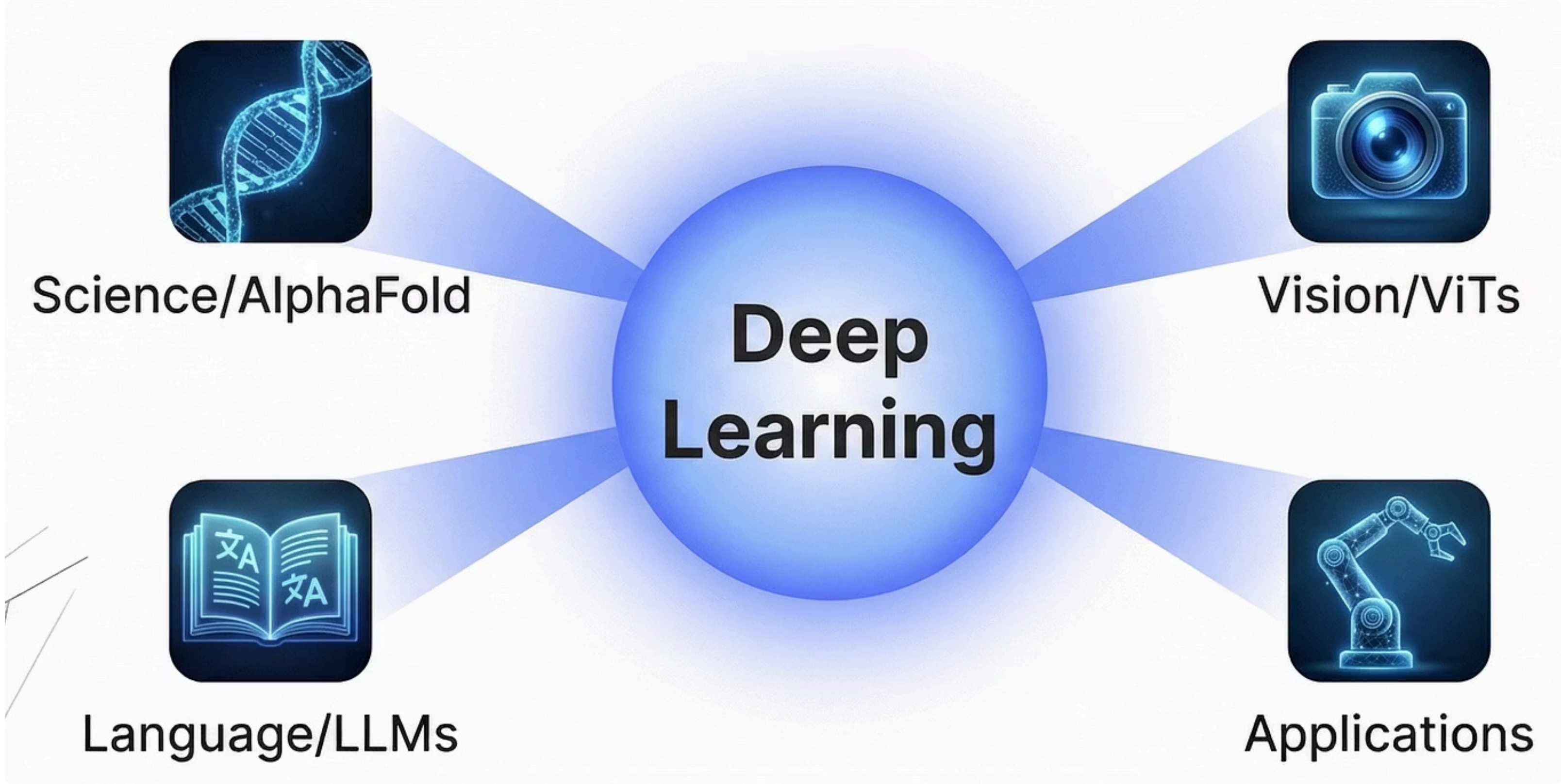
- Monday 11.30 - 13.30
- Check and book here: https://antoninofurnari.github.io/ricevimento/

Member: Image Processing Laboratory (IPLAB)

Research Focus: Embodied AI, Egocentric Vision, Video Understanding

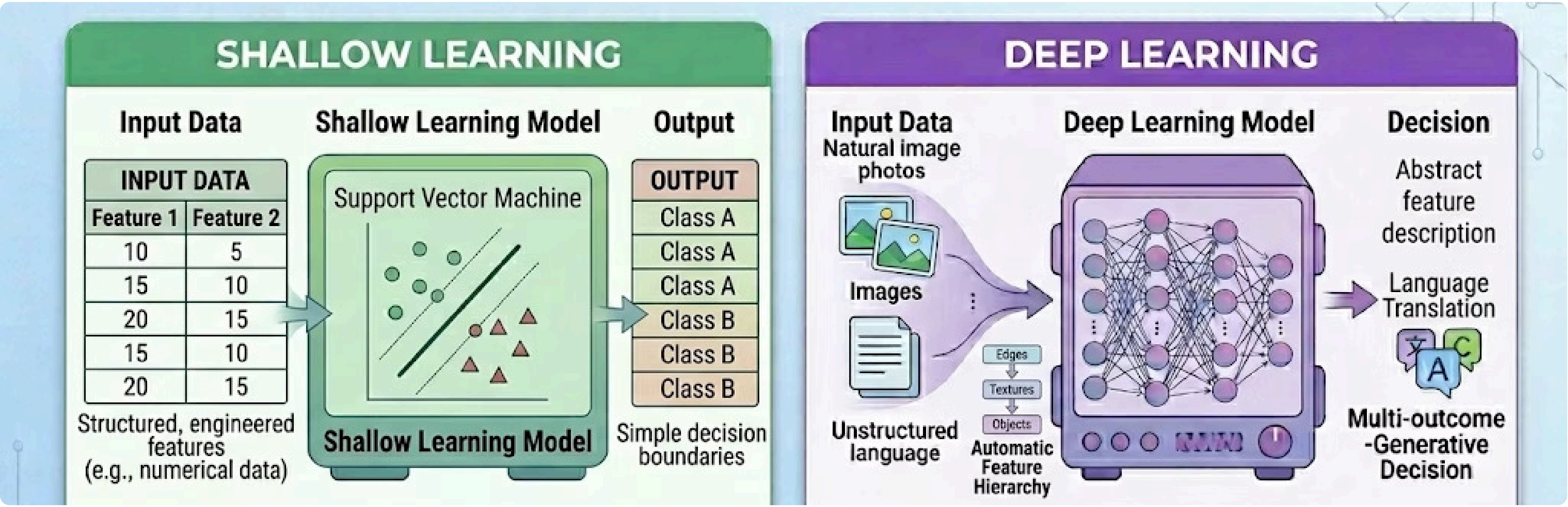
The Core of Modern AI

Deep Learning is no longer a subfield—it's the **operating system** of modern Artificial Intelligence



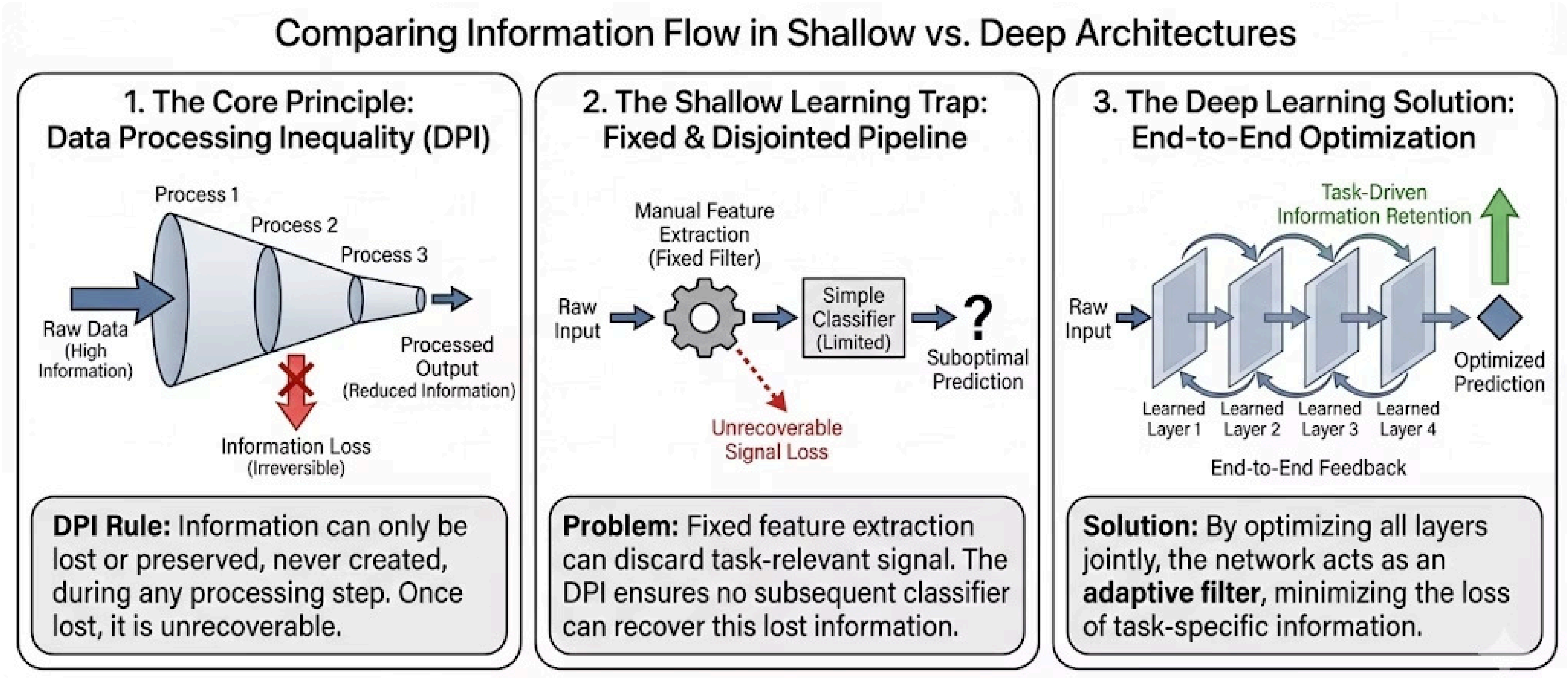
The same mathematical framework—gradient descent on differentiable graphs—now powers text, images, audio, and robotics

The Paradigm Shift
Shallow vs. Deep Learning



This requires a fundamentally different mindset—you're not engineering features, you're engineering learning systems

Shallow vs Deep Learning



Standing on the Shoulders of Giants

1

1958-1969

Perceptrons

Rosenblatt's single-layer networks showed initial promise but hit fundamental limitations, leading to AI's first winter.

2

1986

Backpropagation

Rumelhart, Hinton, and Williams formalized the algorithm for training multi-layer networks, though computational limits prevented scaling.

3

2012

AlexNet Revolution

Krizhevsky's CNN won ImageNet by a massive margin, proving deep learning's viability and sparking the modern era.

4

2017

Attention Mechanism

"Attention Is All You Need" introduced Transformers, enabling unprecedented scaling in language modeling.

5

2020-Present

Foundation Models

Massive pre-trained models like GPT, BERT, and CLIP demonstrate emergent capabilities and few-shot learning.

We are moving from **hand-coded structure** to learned structure

The "Hardware Lottery"

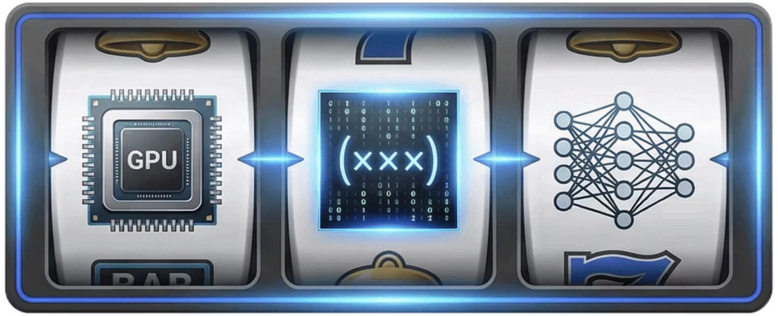
Sara Hooker's Hypothesis

Did Neural Networks win because they are theoretically the best?

Or because they fit perfectly on GPUs?

The Matrix Multiplication Coincidence

- GPUs were built for rendering gaming pixels
- Neural nets are just matrix multiplications
- It was a happy accident of history



THE BITTER LESSON

Why Now?

Rich Sutton's Observation

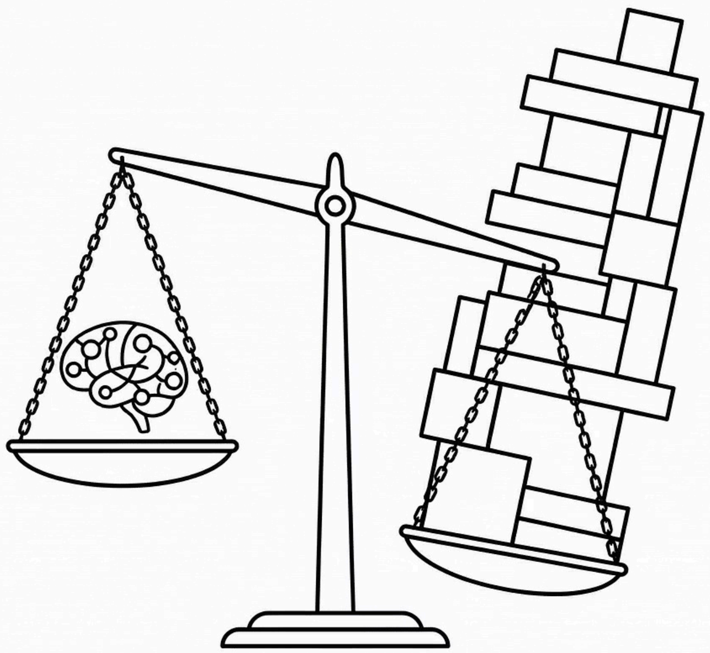
The biggest breakthroughs in AI always come from **general methods that scale with computation**, not from human cleverness

Don't Fight Moore's Law

Every time we hard-coded grammar or chess openings, we lost to a generic learning algorithm scaled up on massive data

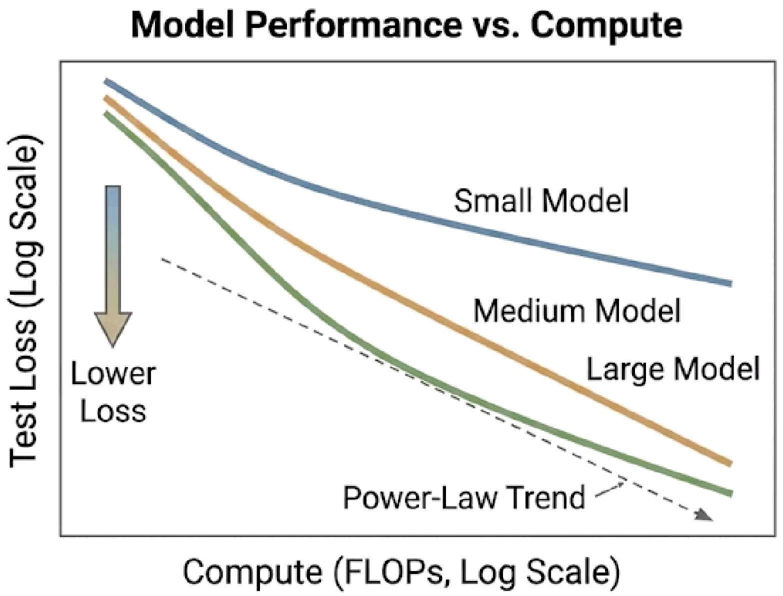
Your Job

Design architectures that *can scale*, not smart hacks that don't



Scaling Laws

THE PHYSICS OF INTELLIGENCE



Performance improves **predictably** with three key factors:

01

More Parameters

Increased model capacity

02

More Data

Richer training examples

03

More Compute

Processing power to train

Where does the architecture / loss function / cleverness of the algorithm stand? It does have a role, especially when data is scarce.

Emergence: At certain scale, models suddenly "learn" capabilities (arithmetic, translation) not explicitly trained.

2024: The Nobel Prize "Anomaly"

A Turning Point in Science

- **Physics Nobel:** Hopfield & Hinton for Artificial Neural Networks
- **Chemistry Nobel:** David Baker, Demis Hassabis, and John Jumper (DeepMind) for AlphaFold

Why This Matters

AI has graduated from "Computer Science" to **"Natural Science"**

> We are no longer just writing software—we are modeling the physical world through learned representations

PHYSICS NOBEL PRIZE

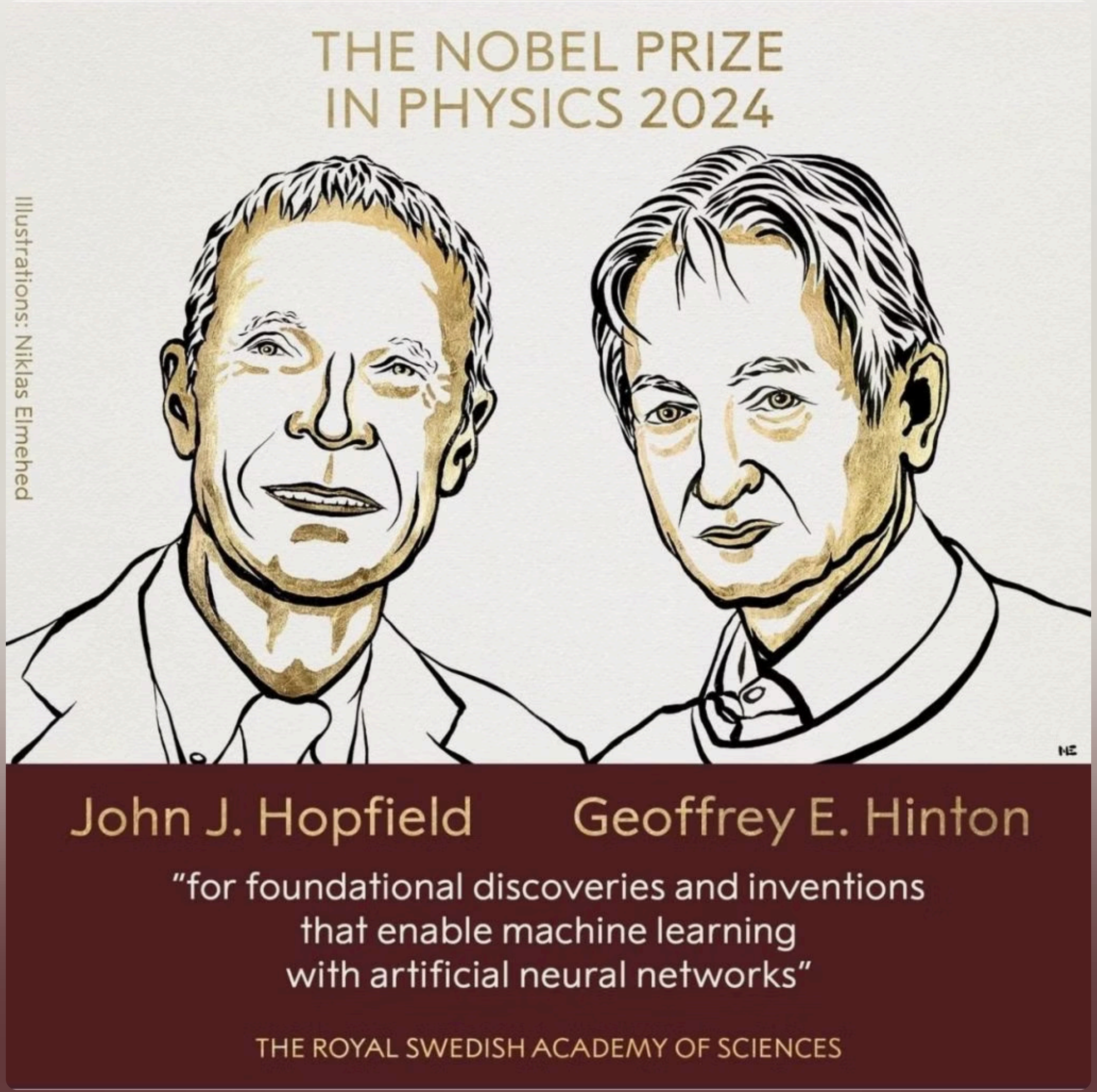
The Physics of Intelligence

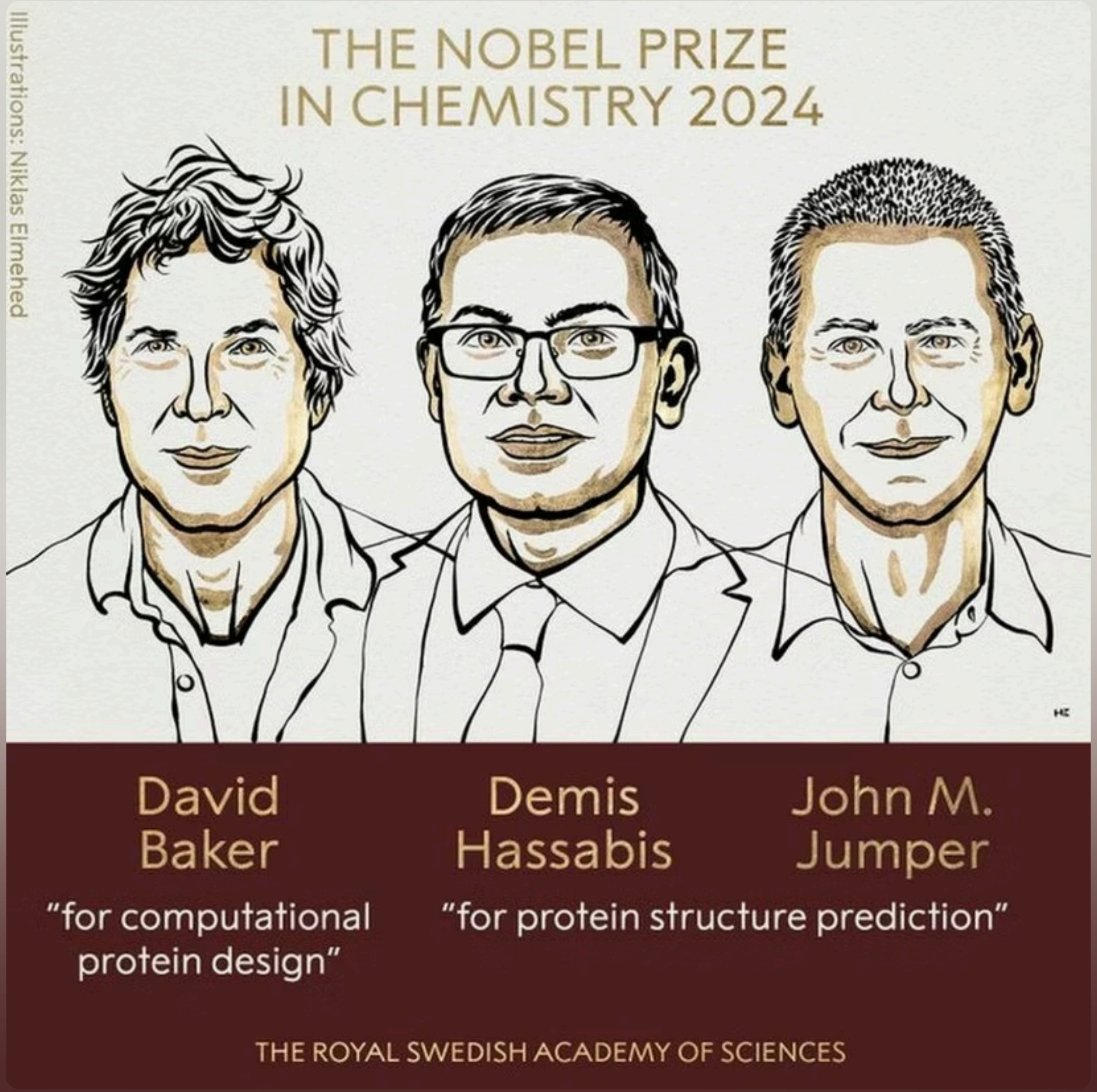
The 2024 Physics Nobel recognized John Hopfield and Geoffrey Hinton's foundational work on neural networks.

Inspired by **statistical physics**, their insights into complex systems revealed how emergent properties arise from interacting components—a core tenet of physics.

This award elevates AI beyond engineering, positioning it as a profound exploration into the fundamental laws of information, learning, and intelligence, deeply rooted in the physical sciences.

Hopfield's work on associative memory networks drew parallels to phase transitions in magnetic materials, while Hinton's contributions laid the groundwork for modern deep learning.





CHEMISTRY NOBEL PRIZE

AI Deciphers Life's Blueprints

The 2024 Nobel Prize in Chemistry honored David Baker for *de novo* protein design, and Demis Hassabis and John Jumper for AlphaFold. This AI system accurately predicts a protein's 3D structure from its amino acid sequence.

This achievement has revolutionized biochemistry, drug discovery, and fundamental life sciences. It highlights deep learning models as indispensable tools for understanding and manipulating life's molecules, making AI a core scientific methodology.

Differentiable Programming

CORE INSIGHT

Deep Learning as Composable Blocks

The "Stitching" Principle

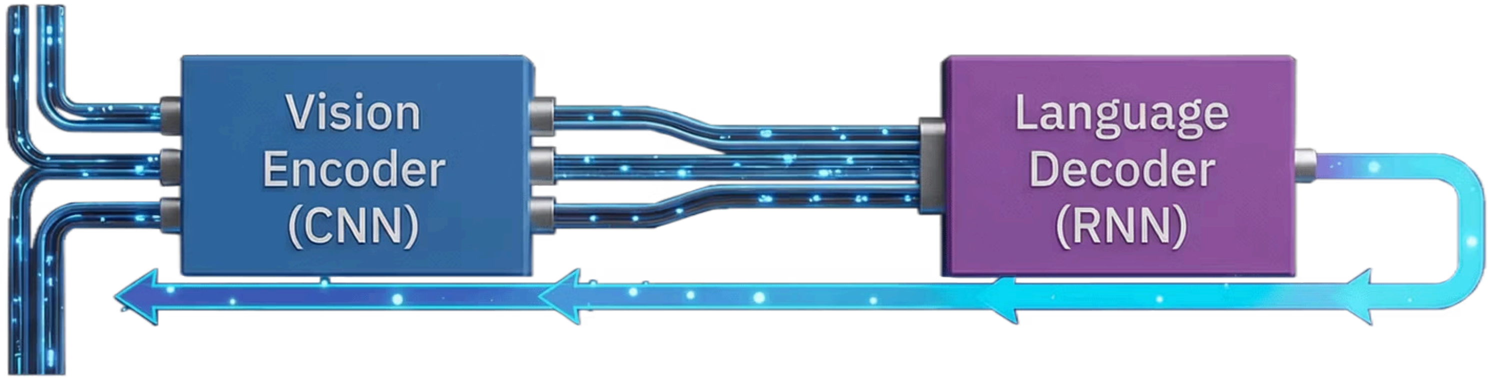
Take a Vision Encoder (CNN) and connect it to a Language Decoder (RNN)

Because both are differentiable, backpropagation flows from text output to pixel input

The "LEGO" Philosophy

This course gives you the bricks

You build the castle



The Manifold Hypothesis

WHY DOES THIS ACTUALLY WORK?

The Curse of Dimensionality

An image is a point in a **3-million dimension space**

Most of that space is noise

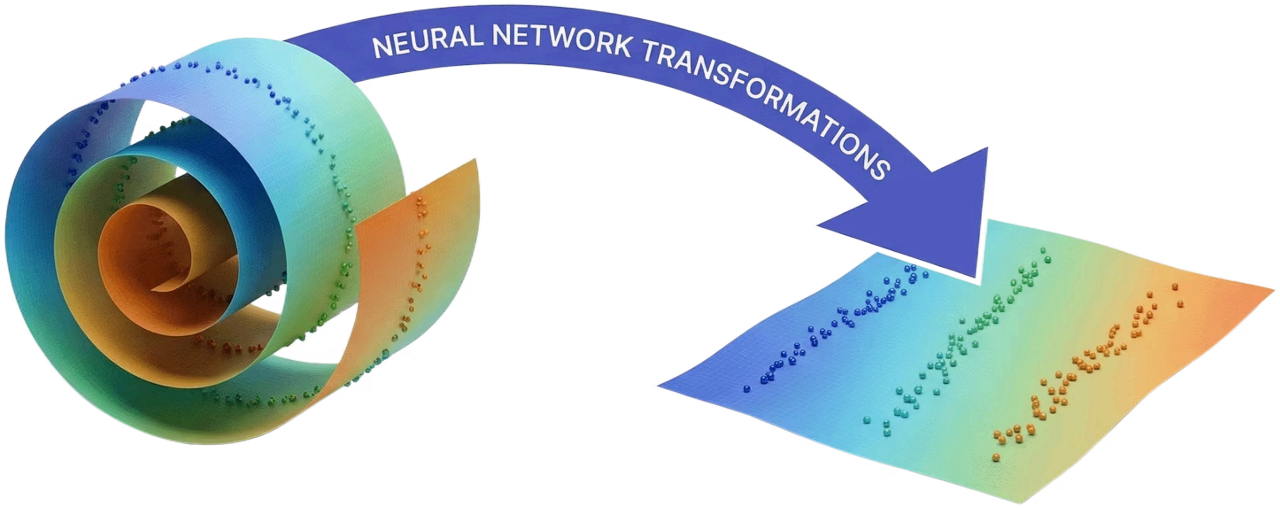
The Hypothesis

Meaningful data (faces, cars) lies on a low-dimensional manifold embedded in that space

What Deep Learning Does

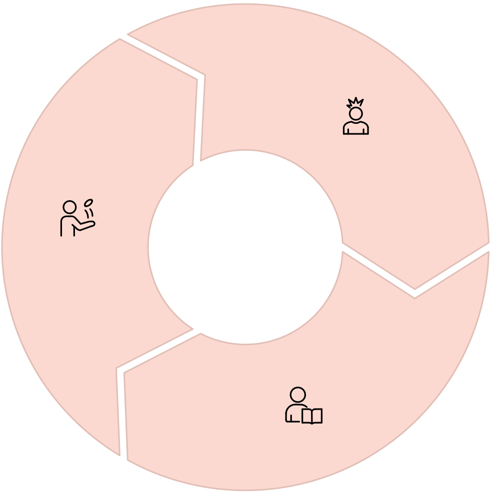
Learns to "unfold" this manifold to make data linearly separable

This is why we need depth—to perform topological transformations



The Era of Experience

Beyond Static Datasets



Agent Acts

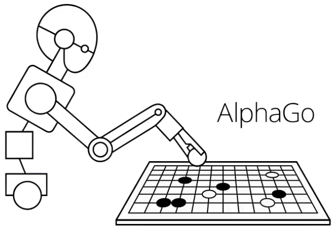
Makes decisions in environment

Environment Reacts

Provides feedback

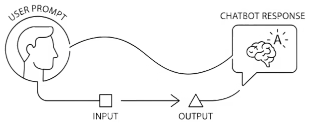
Agent Learns

Updates policy



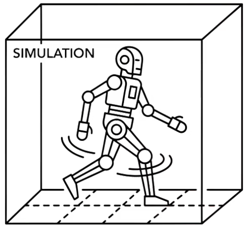
AlphaGo

Learned by playing itself



LLMs (RLHF)

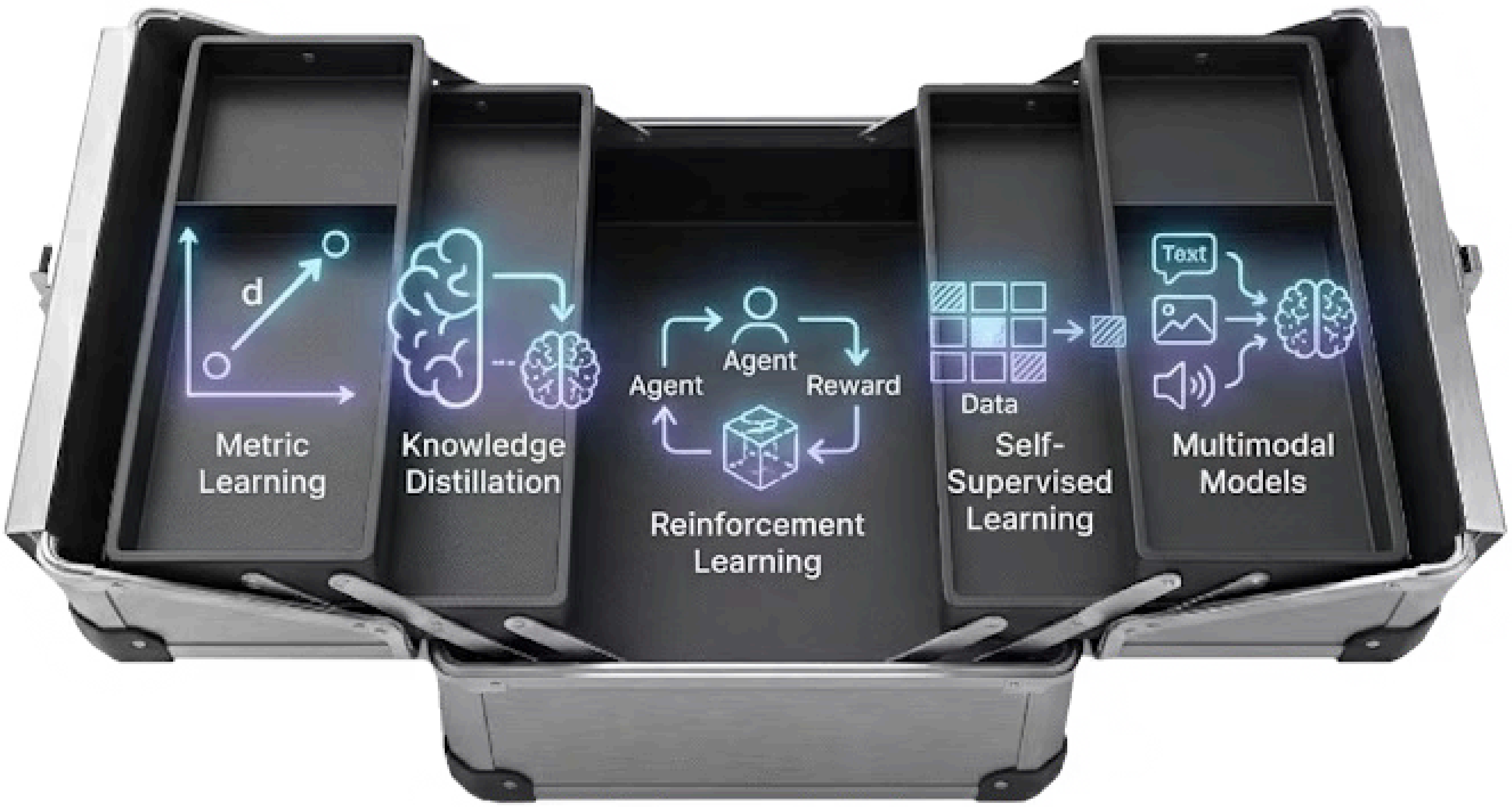
Learned by chatting with humans



Robotics

Learned by trying in simulation

The Modern Toolbox



With Great Power Comes Great Responsibility

The Black Box Problem

Deep models are powerful but opaque

The Scientist's Responsibility

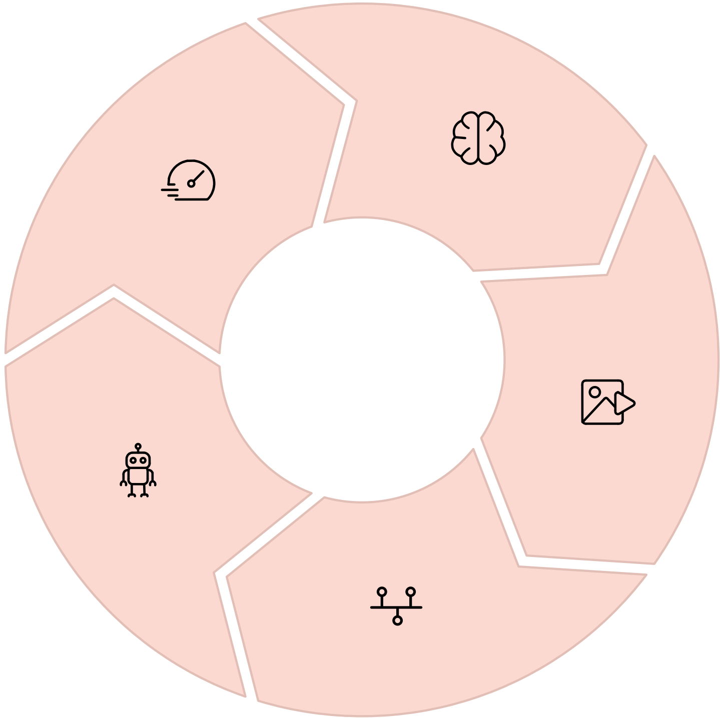
- **Bias:** If data is biased, the manifold is biased
- **Energy:** Large models consume massive energy—is it worth it?
- **Safety:** Adversarial attacks can trick systems easily
- **Agency:** Is the human being still in control?

> **Design Principle:** Ethics is part of the objective function, not an afterthought

Course Structure

THE 5 MODULES

How We Organize The Toolbox



Representation Learning

Learning without labels

Multimodal Models

Integrating senses

Sequential Modeling

Handling time

Agents

Acting in the world

Efficiency & Adaptation

Making it custom, fast and small

MODULE 1

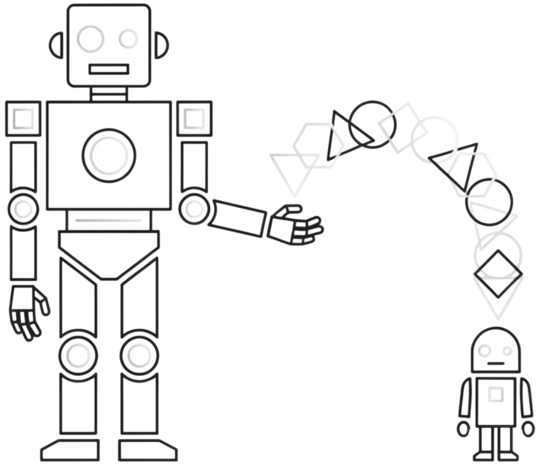
Efficiency & Adaptation

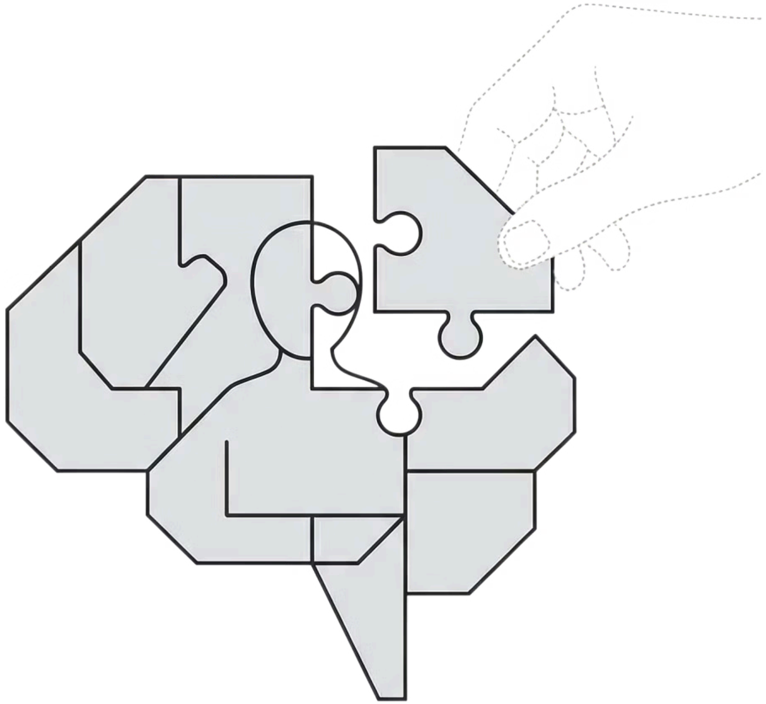
The Challenge

State-of-the-art models are huge. Your phone is small.

Key Techniques

- **Metric Learning:** Recognizing faces with zero-shot learning
- **Domain Adaptation:** Training autonomous driving on simulators (GTA V) to work in reality
- **Distillation:** Compressing a massive "Teacher" into a tiny "Student"





SELF-SUPERVISION

MODULE 2

Representation Learning

The Challenge

Labels are expensive. The internet is full of unlabeled data.

1

Self-Supervised Learning

The engine behind ChatGPT and modern vision

2

Contrastive Learning

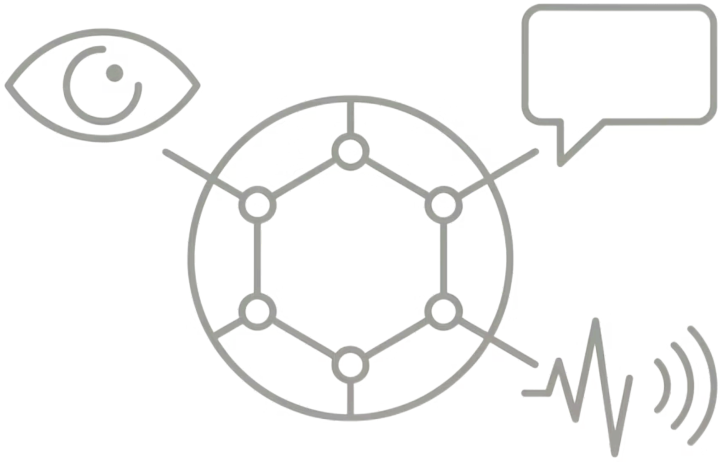
Teaching models that "a cropped dog" is the same as "a rotated dog"

MODULE 3

Multimodal & Video

The Challenge

The world is not just text files—it's a synchronized stream of pixels, audio, and language



1

CLIP

Aligning vector spaces of Text and Images

2

Video Understanding

Adding time dimension (T) to (H × W)

3

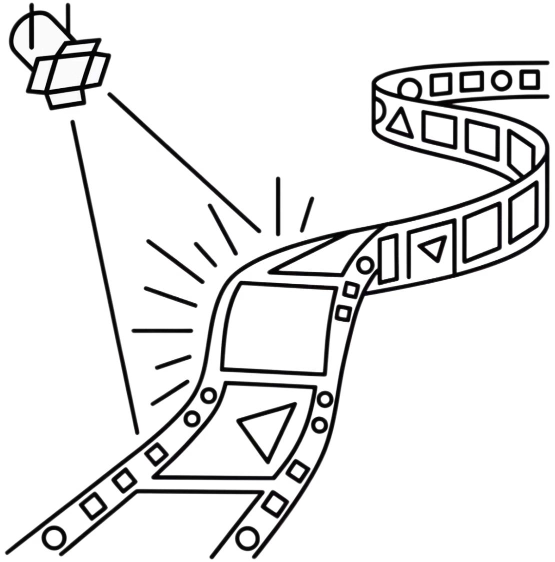
Multimodal Alignment

Stitching vision encoders to language decoders

Sequential & Temporal

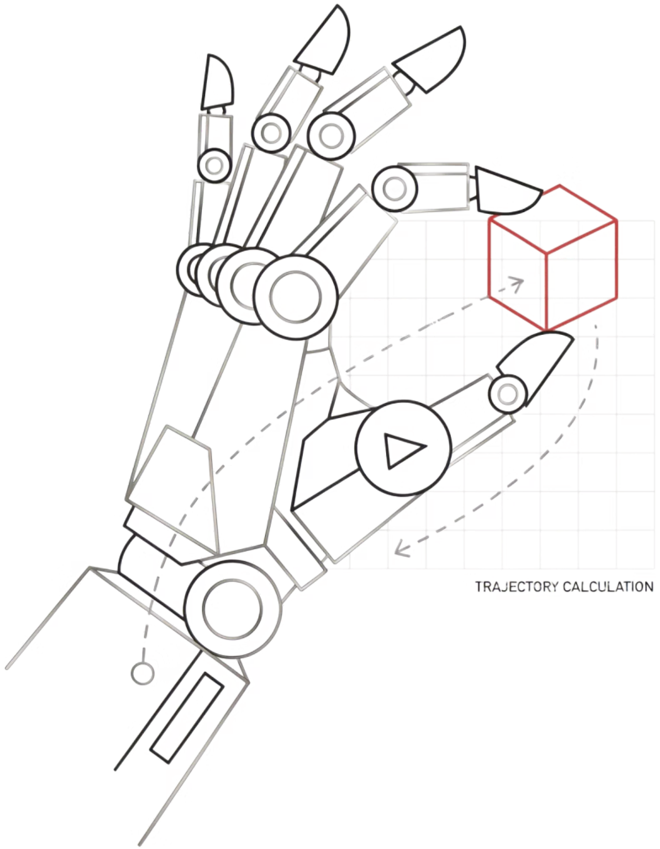
The Challenge: Memory

Context matters. How do we build models that remember?



Key Techniques

- **Online Transformers:** Self-Attention mechanism online (looking at whole context)
- **State Space Models (Mamba):** New efficient alternative for long sequences
- **Application:** Understanding long videos or documents



MODULE 5

The Frontier: Agents

The Challenge

Making **decisions**, not just predictions

Reinforcement Learning

Agents that maximize cumulative reward over time

The Holy Grail

An agent that learns physics and strategy from trial and error

Uni ct

The Scientist Mindset

PART 1: DEBUGGING

Standard Software

If you have a bug → code crashes

Immediate feedback

Deep Learning

If you have a bug → loss goes down... but model learns nothing

Silent failures

Critical Questions

Did you normalize the data? Is the learning rate exploding? Are gradients vanishing?

Skill: You must become a detective of computational graphs

The Scientist Mindset

PART 2: READING PAPERS

No Textbooks

Deep Learning moves too fast. Textbooks are outdated by the time they print.

Primary Sources

Scientific papers remain primary sources to study and understand key concepts.

The Skill

Learning to scan a paper, ignore the fluff, and find the technical details.

Teaching Material & Calendar

Diverse Lecture Formats

Our course combines multiple teaching methodologies to ensure a comprehensive and engaging learning experience, catering to different learning styles and deepening understanding.

Frontal Lectures

Traditional lectures cover foundational theory, core concepts, and algorithmic principles. These sessions build a strong theoretical base essential for advanced topics.

Laboratory Sessions

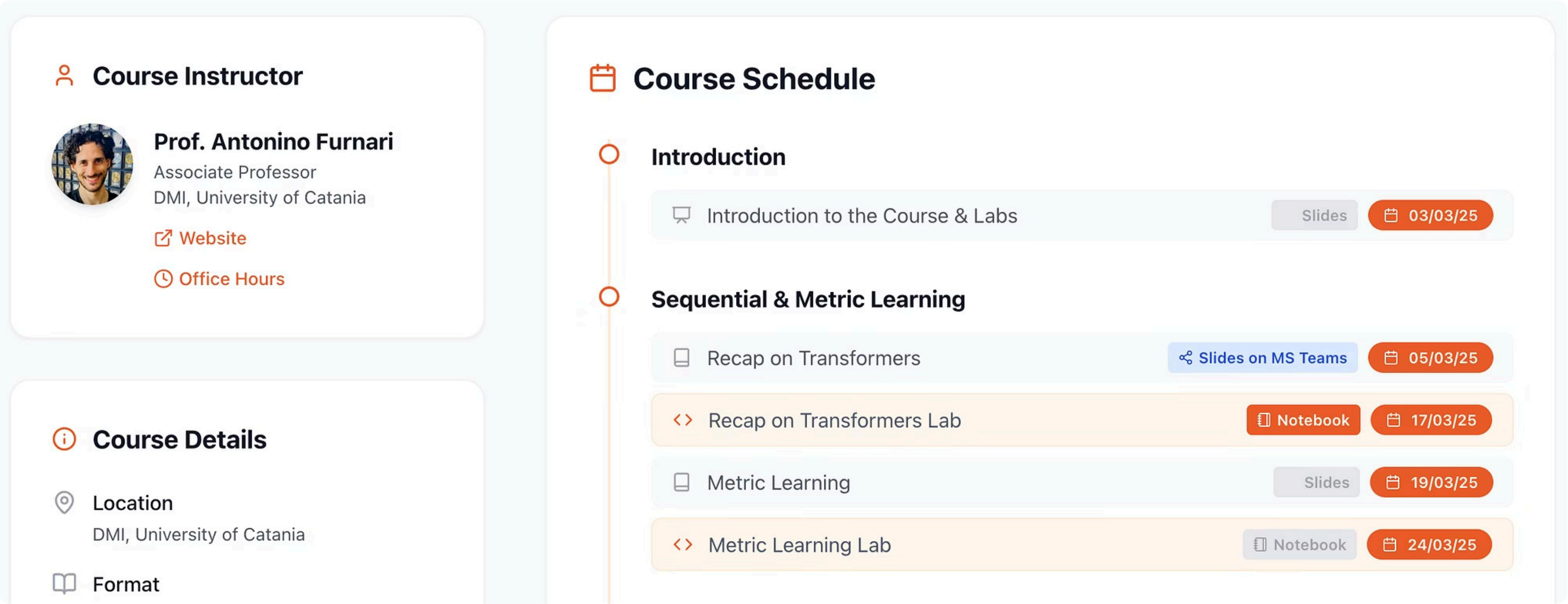
Hands-on coding, practical exercises, and model implementation. Apply theoretical knowledge to solve real-world problems and develop critical practical skills.

Seminars

In-depth discussions on advanced topics, guest speaker presentations, and real-world case studies. Explore cutting-edge research and industry applications in a collaborative setting.

Teaching Material and Calendar

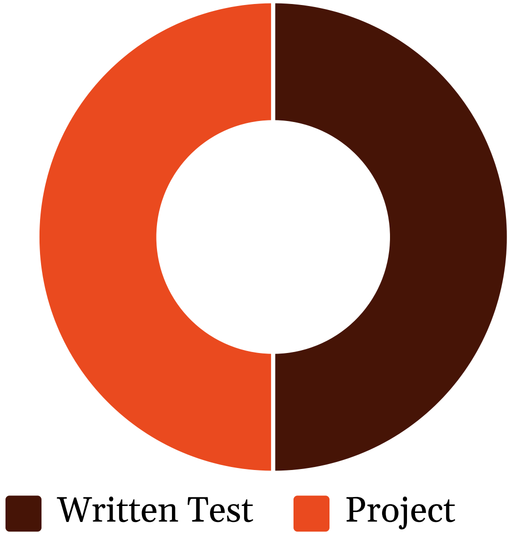
Microsoft Teams (code: cq90ptp)



Website: [Deep Learning: Advanced Models and Methods - UniCT]

Assessment Overview

Philosophy: Understanding Over Memorization



Written Test

Checks your **theoretical foundation**.

Multiple-choice and open ended questions.

Project

Checks your Deep Learning engineering capability.

Assigned by the teacher, developed in groups of 1-3 people and presented live to the teacher for assessment.

The Written Test

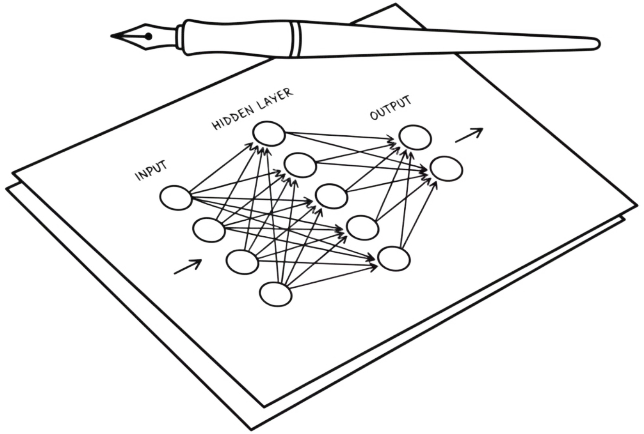
Format

- 30 questions, mixed between multiple-choice and (short) open ended
- 1 hour test / 1 point per question
- I will provide some examples

What We Test

Example: "What problems does Deep Metric Learning tries to solve which are not addressed by standard classification"

Conceptual understanding, not memorization



The Project
Real Experience

01

Receive Project Assignment

Teacher-assigned real-world problem or challenge.

02

Research & Design Solution

Apply techniques seen in lectures to develop a strategy.

03

Implement, Train, Evaluate

Full experimental pipeline to solve the problem.

Format: Groups of 1-3 people | start forming groups now

Presentation: Projects are presented in an oral presentation in the official dates of the exams

Final Mark: (Written test + Project)/2

In-Itinere Tests

1

Mid-Course Test

Covers material from the first half of the course.

15 questions, 15 points.

2

End-of-Course Test

Covers material from the second half of the course.

15 questions, 15 points.

Total score from both tests will be summed.

Project Revisions: Occur on the same days as the in-itinere tests.

Project Presentation: Scheduled for a dedicated day after the course concludes, ensuring completion before the first official exam date.

Laboratory Grading

Practical Sessions

8 hands-on lab sessions, each 2 hours long, conducted in class.

Group Work

Collaborate in groups to complete assigned tasks and exercises.

Submission & Points

Submit your notebook at the end of each session for a Pass/Fail grade. Each passed notebook awards 0.5 points towards your final mark.

Final Mark Calculation

Your overall course mark is computed as: **(In-Itinere Tests + Project) / 2 + Laboratory bonus points**.

Tools of the Trade

YOUR RESEARCH TOOLKIT

PyTorch / Lightning

The lingua franca of AI research

Hugging Face

Repository of pretrained weights and models

WandB

Your lab notebook for tracking experiments

GitHub

Code collaboration and version control

Introduction to Deep Learning Laboratories

Introduction to Deep Learning Laboratories

Course: Deep Learning: Advanced Models and Methods

Objective: Welcome to the course! Today, we aren't just checking our environment; we are establishing the workflow for the entire semester. Deep Learning is as much about engineering discipline as it is about mathematics.

In this lab, we will:

1. **Understand the Notebook Philosophy:** Learn the visual language of these labs.
2. **Verify the Forge:** Ensure PyTorch and CUDA are communicating on Google Colab.
3. **Connect the Logger:** Initialize Weights & Biases (W&B) for experiment tracking.
4. **The Cluster Guide:** A reference guide for moving from Colab to the University Cluster.
5. **First Strike:** Train your first PyTorch Lightning model (MLP on MNIST) and log it.

The Rules of Engagement

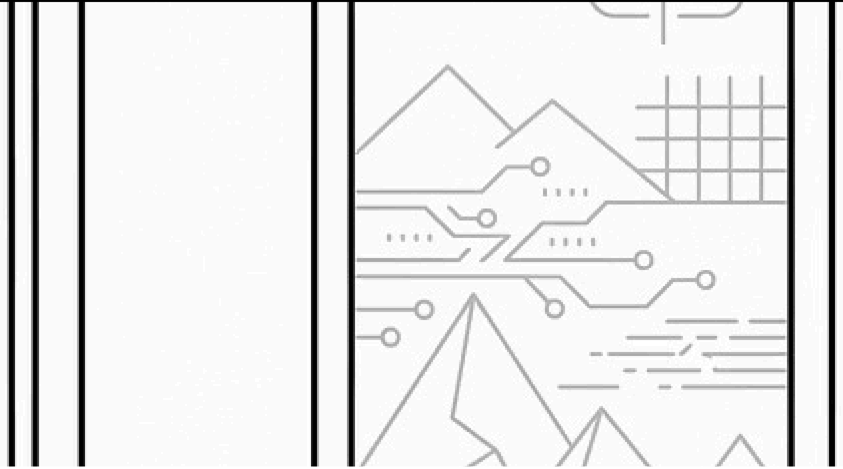
Throughout this course, we will use specific visual cues to guide your learning. Please familiarize yourself with this legend:

> ✅ **TODO** This box provides you instruction on a TODO that should be completed in the code cells below.

> 🛑 **Checkpoint** The following cells provide you instruction on what you should observe if you run a given cell after having implemented your TODOs. This serve as a self-check to assess potential bugs or errors early.

> 📝 **Reflection** You are required to write a short reflection on your results.

> 🏁 **Milestone** This signal that you completed a significant part of the notebook and reached a milestone.



Let's Begin the Journey

Questions?