1. Introduction

Achtung

- I’ve noticed that the course title can be misleading—let me clarify:

  “Deep Learning **Theory**”

- Drop this course, if:
  - This is your first deep learning course — Take “Deep Learning” instead
  - You want to understand why deep learning works — “Science of DL” usually helps more
  - You want to learn how convolution, self-attention works — We talk about MLP all the day
  - You don’t know how to prove mathematical claims — That’s all we do in this course

Achtung

- As an example, you'll be proving claims like this the whole semester:

Theorem

Let continuous $g : \mathbb{R}^d \to \mathbb{R}$ and $\epsilon > 0$ be given. Choose $\delta > 0$ so that

$$\|x - x'\|_\infty \leq \delta \longrightarrow |\, g(x) - g(x') \,| \leq \epsilon$$

holds. Then, there exists a 3-layer network $f(\,\cdot\,)$ with $\Omega(1/\delta^d)$ ReLU neurons, with

$$\int_{[0,1]^d} |\, f(x) - g(x) \,| \, \mathrm{d}x \leq 2\epsilon$$

- If you don't like what we are doing, let's not waste our time ;)
    - Busy, aren't we?

Then why theory?

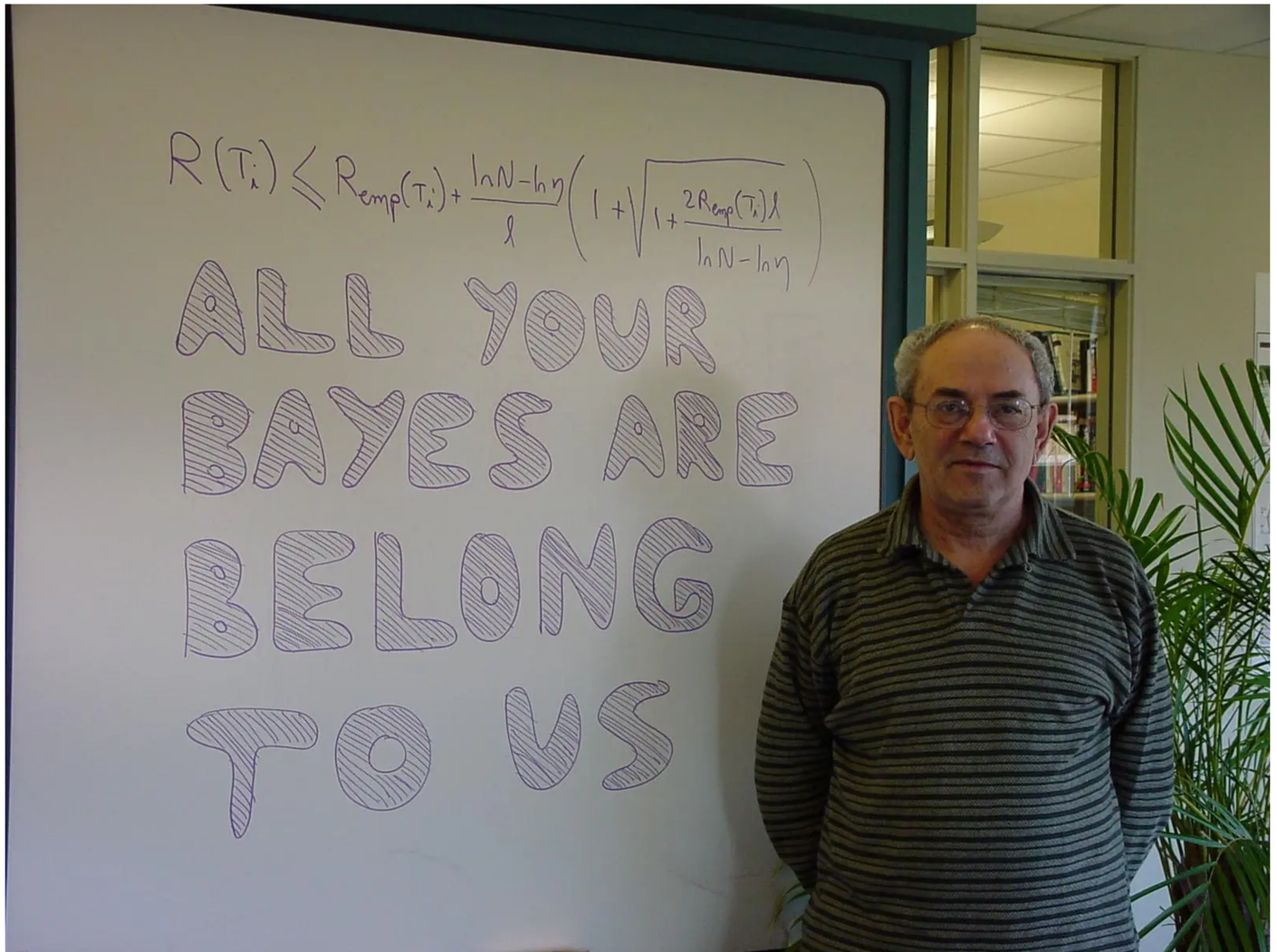
- A few reasons:
  - Theories give you **perspectives**
    - e.g., “information theory” views signal processing as a matter of preserving information.
  - Theories equip you **formal logic**
    - formalizing your intuition to a precise statement
  - Experiments are getting more and more **expensive** nowadays
    - to quote Misha Gromov: “math is a part of physics where experiments are cheap”
    - need a correct and useful mathematical model, of course

Then why theory?

- By taking theory courses, I expect you to do the followings (from the easiest to the hardest)
  - Decorate your ML algorithms paper with theorems.
  - Motivate new ML algorithms from theories
  - Demystify your empirical observations
  - Deduce new knowledge
  - Have fun

In this course

- We are going to take a framework, that is known as **statistical learning theory**
  - More frequentist than Bayesian

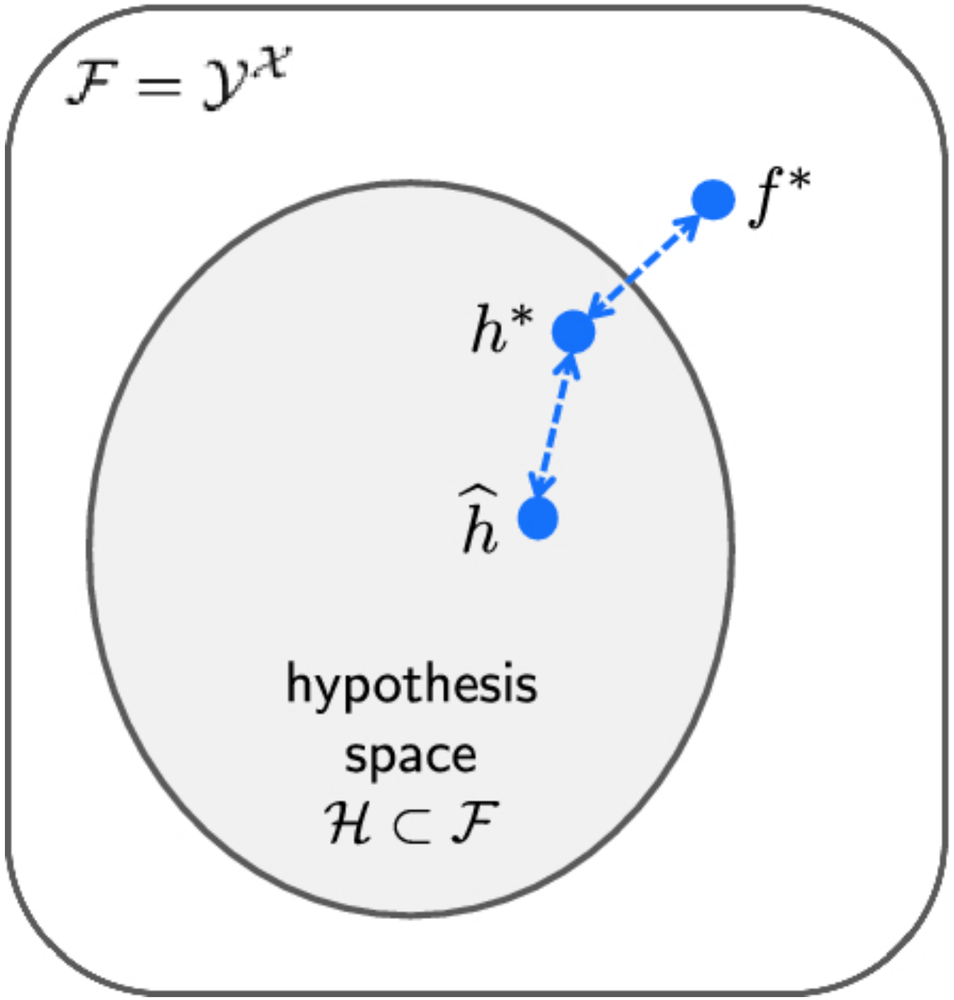


In this course

- Treat neural networks as **a parameterized set of functions**
  - also known as hypothesis space

- We ask the questions like:
  - How expressive is this set?
  - How well can we optimize over this set?
  - How generalizable are the optimized solutions?

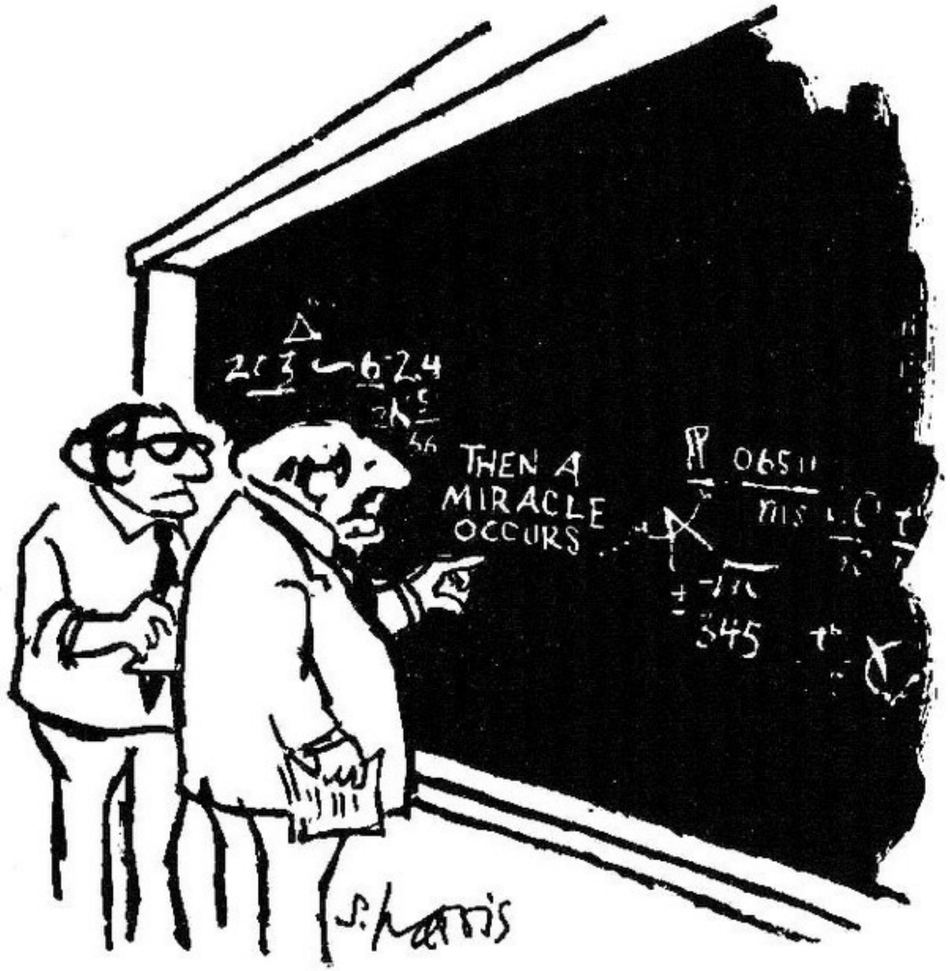
- We care about theoretical guarantees that we can (or cannot) do certain things
  - e.g., under condition X, an event Y will happen, with probability at least Z



Caveat

- We focus on classical theories, which have failed to fully explain "why deep learning works":
  - However…
    - we don't have a better option ;(
    - modern theories still build on these

- Proofs will be somewhat informal
  - Focus on delivering intuitions
  - In fact, typical in actual papers
  - See the textbook for the full proof
    - Or even better, do it yourself



"I think you should be more explicit here in step two."

Administrivia

Team

- **Instructor**
  - Jaeho Lee 이재호
  - jaeho.lee@postech.ac.kr
  - Lectures, Q&A

- **Teaching Assistant**
  - Byung-Ki Kwon 권병기
  - byungki.kwon@postech.ac.kr
  - Assignments, Attendance

Locations & Hours

- **Lectures**
  - Engineering Building 2, Hall 109
- **Office Hours**
  - GoAround Coffee @ RIST
  - Mondays, 5PM — 6PM
- **Materials**
  - https://jaeho-lee.github.io — Lecture notes
  - PLMS — Assignments and videos

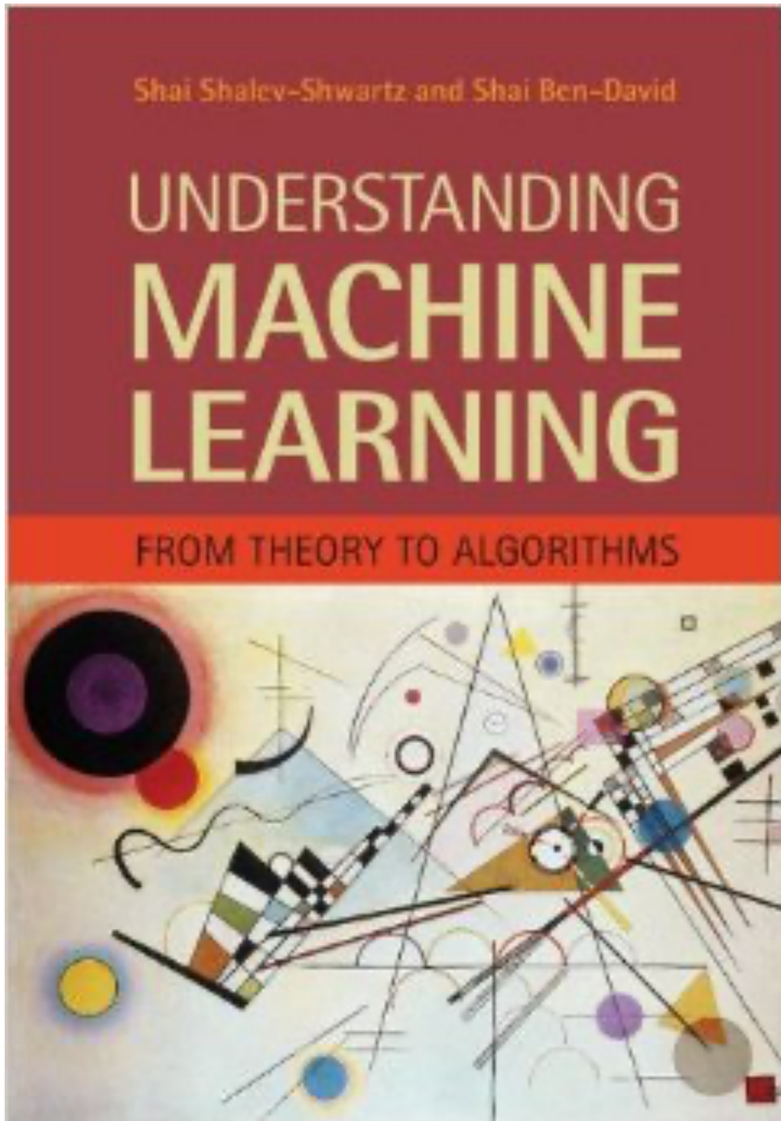
Textbooks

Main

- MJT's lecture note: https://mjt.cs.illinois.edu/dlt/index.pdf
    - We'll use the old version
    - There is a new version too, which is incomplete (sadly)

Sub

- SSS & SBD (GOAT)
- https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/



Grading

- **Attendance (10%)**
- **Assignments (30%)**
  - Planning to give you 3 homework assignments
- **Mid-term (30%)**
  - New this semester
- **Final Project (30%)**
  - Prove a new result, relevant to your research (Best)
  - Read a recent paper and write a 3-page report to summarize it
    - Describe the main result, and prove it in a self-contained manner
    - You may make simplifying assumptions, for the sake of clarity
    - The paper should have been published in 2024 or 2025

Attendance

- Please use the **electronic attendance** system

- You'll get an **F**, if you miss **more than 3** classes.
  - What does not count:
    - "I was there, but I forgot to …"
    - "The system did not work …" (unless others experienced the same issue)
    - "I had a group event…"
  - What counts, conditionally:
    - International Conference — if you are the first author
    - Sick — if you can prove it
    - Family emergency — if you can prove it

Prerequisites

- I'll assume that you know:
  - How to formally prove things
  - Deep learning
  - Real analysis (bits)
  - Probability and random processes (bits)

Honor codes

- You'll get an **F**, if:
  - Sharing solutions
  - Copying solutions
  - "Collaborating" with your friends for homework
    - Use me or TA instead
  - Using GPT for proofs
  - Again, miss more than 3 classes