ECE 491B: Introduction to large-scale AI systems

This course will focus on Large Language Models (LLMs): their internal mechanisms, their development, their deployment, evaluation, and their integration with tools. We will go over NLP and ML fundamentals, discuss what shifts of paradigm prompted the appearance of large-scale AI, gain experience working with cloud computing infrastructure, and get introduced to state-of-the-art research. We will highlight interdisciplinary applications as well as socioeconomic and ethical issues associated with AI and LLM in particular.

Aim: for students to gain the engineering capability to reproduce empirical results from LLM-related AI papers.

Pros:
Accessibility: Live lectures + recordings + slides
Practicality: Hands-on assignments + Final project

Cons:
None really

Pre-requisites: linear algebra + solid coding background (Python, PyTorch, bash, git)

Keywords: Self-supervised learning, transformers, distributed training

More info:

Syllabus

Website