Economics 2355: Unleashing Novel Data at Scale

Spring 2024

Many important economic questions remain unanswered, in substantial part because the data required to examine them has traditionally been inaccessible. Information that could elucidate important questions is scattered throughout text, or contained in scans, photographs, satellite imagery, videos, or hard copy documents. This course will provide an introduction to ***deep learning-based (AI)*** methods for converting **unstructured data** into computable information at scale. It will cover workhorse architectures and methods and examine a range of vision, NLP, and multimodal applications. This course *does not* cover machine learning methods for structured data.

This course differs from a computer science course in its intensive focus on methods that are most likely to be useful for processing unstructured data in economics. Our applications are often different in fundamentally important ways from those in the computer science literature, and a key theme of the course is understanding the aspects of a deep learning model that make it most suited to academic applications. The deep learning literature is very large and moves at a fast pace. It is my hope that by consolidating this information into a course, it will save others a lot of the inefficiency that comes with navigating a massive and fast-moving literature. This is a very nascent discipline, and the course covers the most cutting-edge methods, with a large number of the readings on the syllabus from 2022. The course has been completely redesigned since when it was previously offered in 2020.

Instructor:

Melissa Dell

M-24, Littauer Center

Department of Economics

melissadell@fas.harvard.edu

Teaching Assistants:

Tom Bryan

tom_bryan@fas.harvard.edu

CGIS Knafel K317

Office hours: Wednesdays 2-4:30pm, in-person at CGIS South S314 and on Zoom (https://harvard.zoom.us/my/bryanptom)

SyllabusEc2355_2024_2.pdf

First Day Slides:

Lecture1_2024.pdf