

EC9C3 - Topics in Industrial Organisation and Data Science

Mirko Draca and Ao Wang

October 2, 2024

1 Introduction

The goal of this course is to get you thinking about sophisticated research design. What actually goes into writing an interesting paper? This is hard to formalise but one answer is that it is about answering creative questions in new, convincing ways.

A big part of the ‘new ways’ aspect is the thoughtful and clever use of empirical tools. In this course we will introduce you to a range of methods that have gained prominence recently from the areas of machine learning (aka ‘data science’) and structural IO. But research should not be about using techniques for the sake of it! So we will also centrally focus on how these methods have been used in good papers.

2 Course Sections

2.1 Topics in Data Science (Draca)

This part of the course will be structured as follows:

- **NLP Building Blocks:** term-document matrices, TFIDF and cosine distance.
- **Model Selection and Regularization:** Tools for handling a big covariate X space, such as subset selection, Lasso and Ridge Regression. How have these tools been adapted and used in economic applications?
- **Supervised Learning:** We’ll look at three classification techniques (i) Naive Bayes, (ii) Regression Trees, and (iii) Support Vector Machines.
- **Unsupervised Learning:** How can we dig out ‘hidden structure’ in some some space of X variables? We’ll discuss Principal Components, K-means Clustering and Topic Models.
- **Introduction to Research Design:** A session to get you looking at some clever and creative papers. Nothing that is actually directly related to data science methods but will inform how you come up with papers (if we have time).

2.2 Topics in Industrial Organisation (Wang)

The usual structure for this section is:

- Introduction to Structural Models in (Empirical) Industrial Organization.
- Models of Demand for differentiated products: methods with macro and micro data.
 - AIDS, Semi-parametric approach.
 - BLP models.
 - Advanced topics (e.g., multiple purchases, complementarity in demand)
- Models of Supply I: pricing and market power
 - Pricing game, merger analysis, simulate an excise tax.
- Model of Supply II: Bargaining

Most lectures will be accompanied by an illustration of applications in recently published papers.

3 Assessment

Assessment will be via two extended homework assignments, weighted as 50% each. They relate separately to each part of the course, with the following due dates:

- Monday, 11 November 2024
- Monday, 16 December 2024

(Note: Above dates are subject to final confirmation).

4 Rooms

- Tuesday 9-11am - Room S2.81A (Social Science).
- Wednesday 9-11am - Room S2.81A (Social Science).

Balance of face-to-face and virtual delivery to be discussed.