



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Economics For Data Science M

2021-2-F8204B020

Obiettivi formativi

Al termine del corso, gli studenti avranno imparato come:

- Costruire un modello economico per identificare e stimare un impatto causale.
- Confrontare e valutare diverse strategie di econometria applicata per stimare un impatto causale.
- Usare big data e machine learning per la stima di un impatto causale.
- Capire vantaggi e valore aggiunto dell'utilizzo di big data per ricerca applicata nelle scienze sociali.

Al termine del corso, gli studenti sapranno:

- Utilizzare le tecniche di program evaluation per rispondere a domande di ricerca che pongono problemi di policy-relevance.
- Utilizzare in modo efficace i big data per rispondere a importanti domande di ricerca economica.

Contenuti sintetici

Questo corso introduce il field emergente che nasce dalla fusione di Economia e Data Science per rispondere a domande di policy-relevance. L'obiettivo principale del corso è discutere le più importanti tecniche di machine

learning e di econometria applicata per stimare un impatto causale, e i vantaggi dell'utilizzo di big data per rispondere a importanti domande di ricerca in diverse applicazioni.

Discuteremo i seguenti quattro argomenti principali:

- 1) Validità del Modello e Inferenza Causale.
- 2) Machine Learning e Inferenza Causale.
- 3) Modelli Strutturali, Esperimenti e Machine Learning.
- 4) Applicazioni Empiriche che usano Big Data.

Programma esteso

Topic 1: Model's Validity and Causal Inference.

- Internal and external validity
- Big data: new frontiers for economic analysis

Topic 2: Machine Learning and Causal Inference.

- Program evaluation and the "missing" counterfactuals
- Causal inference and machine learning

Topic 3: Structural Models, Experiments and Machine Learning

- Developing, estimating and using structural models
- Comparing experiments, structural models and machine learning methods for causal inference

Topic 4: Empirical Applications Using Big Data

- Students' presentations: present your own work, one of the papers from a list of suggested papers that will be provided, or a paper of your choice that uses machine learning methods, possibly replicating the results of the paper you choose to present.

Prerequisiti

Principi di econometria applicata e metodi quantitativi di statistica applicata.

Metodi didattici

Lezioni e progetti degli studenti (individuali o di gruppo).

Modalità di verifica dell'apprendimento

Presentazioni e progetti di applicazione dei modelli e metodi economici ai dati.

Testi di riferimento

Libri di testo: per questo corso non c'è un libro di testo di riferimento. Di seguito un elenco di alcuni testi di riferimento per gli argomenti di econometria trattati nel corso. Tutti i libri di testo di seguito riportati sono disponibili in formato e-book tranne Wooldridge (2020) che è disponibile presso la Biblioteca di Ateneo sede Centrale e Sede di Scienze.

Avanzato:

- W. H. Greene. *Econometric Analysis*, 5th Edition, Prentice Hall International, 2003.

Semplice/meno tecnico:

- J. Wooldridge. *Introductory Econometrics: A Modern Approach*, 7th Edition, Cengage Learning, 2020. (for IV and 2 stage least squares)
- Stock and Watson, *Introduction to Econometrics*, 3rd Edition. (Basic statistics and regression analysis; companion website with datasets and files to replicate empirical results: https://wps.pearsoned.com/aw_stock_ie_3/178/45691/11696965.cw/index.html)
- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009. (Excellent for concept of causality, experiments, diff-in-diff, and RD)

Articoli e capitoli di libro: la discussione di ciascuno dei quattro argomenti di cui discuteremo nel corso farà riferimento ai seguenti principali articoli:

Topic 1: Model's Validity and Causal Inference.

- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009, Chapter 1 pages 3-8.
- Athey, Susan, 2017. "Beyond prediction: Using big data for policy problems", *Science*, 355, 483-485.

- Einav L., and J. Levin. 2014. "Economics in the Age of Big Data", *Science*, Vol 346, Issue 6210: 1243089.
- Kleinberg, John, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. "Prediction Policy Problems" *American Economic Review*, 105(5): 491–95.
- Kleinberg, John, Jens Ludwig, and Sendhil Mullainathan. 2016. "A Guide to Solving Social Problems with Machine Learning", *Harvard Business Review*.
- Shmueli, G. 2010. "To Explain or to Predict?", *Statistical Science*, Vol. 25, No. 3, 289-310.
- Varian, H. 2014. "Big Data: New Tricks for Econometrics", *Journal of Economic Perspectives* 28, 3-28

Topic 2: Machine Learning and Causal Inference.

- Athey S. and G. Imbens. 2015. "Machine Learning Methods in Economics and Econometrics", *American Economic Review: Papers and Proceedings*, 105(5): 476-480.
- Varian, H. (2016) "Causal Inference in Economics and Marketing", PNAS, Vol. 113, No. 27, pages 7310-7315.
- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009, Chapter 2 pages 11-24, Chapter 6 pages 251-267.

Topic 3: Structural Models, Experiments and Machine Learning.

- Attanasio O., Costas Meghir and Ana Santiago. 2011. "Education choices in Mexico: using a structural model and a randomised experiment to evaluate Progresa", *Review of Economic Studies*.
- Todd, Petra, and Kenneth, I. Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioural Model of Child Schooling and Fertility", *American Economic Review*, Vol. 96, Issue 5, pages 1384-1417.

Topic 4: Empirical Applications Using Big Data.

References: a list of papers will be provided.

Periodo di erogazione dell'insegnamento

Secondo semestre.

Lingua di insegnamento

Inglese.
