Syllabus

Course: StatisticalLearningforEconomicsandFinance
Professor:

SUMMER 2024

This course describes the main techniques in statistical learning and their applications to economics and

COURSE OUTLINE

finance. We start with the definition of statistical learning relative to machine learning and high-dimensional econometrics, and then discuss big data in economics and finance. We next go through the details of regularization methods, describing not only their corresponding intuition, but also the challenges of carrying out statistical inference. As regularization methods typically involve automatic model selection, we briefly discuss how to do inference after model selection. Next, we introduce statistical learning algorithms for nonlinear models, including tree-based regressions and artificial neural networks, as well as for the analysis of text as data. We conclude with a brief discussion about causality.

METHODOLOGY

Students must read references before lectures, in order to foster discussion.

PROGRAM

1. Introduction: Definitions and review of applications
2. Linear models: Regularization methods, sparsity assumptions, inference after model selection
3. Nonlinear models: Universal approximation by trees/forests/neural networks, sieves estimation
4. Text as data: NLP techniques and applications in economics and finance
5. Causality: DAG and, if time allows, causal trees

BIBLIOGRAPHY

Chan & Matyas (Econometrics with ML, 2022), Masini, Medeiros & Mendes (Journal of Economic Surveys 2022), Athey (NBER 2019), and Chetty (Harvard course 2019).

Bryzgalova (2015), Bryzgalova & Julliard (2019), Bryzgalova, Pelger & Zhu (2019), Callot, Kock & Medeiros (JAE 2017), Chen, Pelger & Zhu (2019), Chinco, Clark-Joseph & Ye (JF 2019), Engle, Ledoit & Wolf (JBES 2019), Feng, Giglio & Xiu (2017), Feng, Polson & Xu (2018), Fernandes, Medeiros & Scharth (JBF 2014), Fernandes & Vieira (JECD 2019), Gu, Kelly & Xiu (RFS 2019), Ledoit & Wolf (JEF 2003, JMA 2004, JPM 2004, AoS 2012, RFS 2017), Medeiros, Vasconcelos, Veiga & Zilberman (JBES 2019), Olivares-Nadal & DeMiguel (OR 2008).

GRADING

Marks rest on participation (20%) and on the students' critical assessment of a research paper that applies machine/statistical learning methods to either economics or finance (80%). In particular, for the latter, students must hand in referee and replication reports and do a 15-minute presentation.

PROFESSOR – EMAILS