# Machine Learning in Economics (ECON3389) - Syllabus

Fall 2018

**Lectures** Mondays and Wednesdays 8:30-9:45 at O'neill Library 247

**Instructors**
Prof. Stefan Hoderlein
Joe Cooprider (PhD Candidate)

**E-mail** Please include "ECON3389" in the subject field
Stefan Hoderlein stefan.hoderlein@bc.edu
Joe Cooprider coopridj@bc.edu

**Office Hours**
Stefan Hoderlein Mon, Wed 9:45 AM-10:30AM
Joe Cooprider Tue 1:30PM-3:00PM

**Description** Large-scale data sets ("big data") have become ubiquitous across many applied areas. The goal of this course is to provide an introduction to methods that allow to deal with this situation. We focus on statistical learning techniques and high-dimensional statistics, and show how they can be applied in economics and business administration. Students will learn how to program statistical methods in PYTHON, as well as how and when to use the common libraries in these languages.

**Prerequisites** For students majoring in Economics, we require Economic Statistics (ECON1151) and Econometric Methods (ECON2228). Students coming from different departments are welcome, but should have similar

command of statistical methods. A solid knowledge of differential calculus at the level of MATH1102 (the "preferred" co-requisite for ECON2228) is highly recommended, as well as fundamentals of linear algebra (matrix notation, multiplication, inverses, determinants). Prior knowledge of programming is *not* a prerequisite, but student should be willing to learn it.

**Homework** There will be bi-weekly homework assignments that will include programming, mathematical problems and applications on real and simulated data sets.

**Textbooks**

1. *Main textbook:* James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. New York: Springer. Available for free at: http://www-bcf.usc.edu/∼gareth/ISL/

2. *Other references:*

   (a) Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.

   (b) Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer, 27(2), 83-85. Available for free at: http://statweb.stanford.edu/∼tibs/ElemStatLearn/

   (c) Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

   (d) Giraud, C. (2014). Introduction to high-dimensional statistics. CRC Press

   (e) Hastie, T., Tibshirani, R., Wainwright, M. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations. CRC Press.

**Course Outline** This is the first time this course will be taught, so please be advised that the material may change during the semester.

1. Introduction: statistical models, loss functions, optimization
2. Review of multivariate linear regression
3. Beyond linear regression: nonlinear regression, polynomial regression
4. Learning theory: model selection, bias-variance trade-off, overfitting and underfitting, penalization.
5. Regularization: Ridge regression
6. Sparsity: LASSO
7. Ensemble methods: Random forests, Boosting
8. Classification: Logistic regression, Neural net/Deep Learning
9. Dimensionality reduction: Nearest-neighbors clustering, Principal component analysis (PCA).

**Grading**

- Homeworks: 30% (best 6 out of 7)
- Midterm: 30%
- Final project: 40%

| | | |
|---|---|---|
| | Aug 27th (Mon) | First day of classes |
| | | Homework 0 (math review) handed out |
| | Sep 5th (Wed) | Homework 0 (math review) due |
| | | Homework 1 handed out |
| | Sep 19th (Wed) | Homework 1 due |
| | | Homework 2 handed out |
| | Oct 3rd (Wed) | Homework 2 due |
| | | Homework 3 handed out |
| **Schedule** | Oct 17th (Wed) | Homework 3 due |
| | | Homework 4 handed out |
| | Oct 31st (Wed) | Homework 4 due |
| | Nov 7th (Wed) | **Midterm** |
| | | Homework 5 handed out |
| | Nov 26th (Mon) | Homework 5 due |
| | | Homework 6 handed out |
| | Dec 5th (Wed) | Homework 6 due |
| | Dec 19th (Wed) | **Final Project due** |

Please note that there will not be rescheduled or make-up examinations. Homework assignments will not be accepted past their due dates. You must demonstrate your reasoning and show all calculations to receive full grade.

**Academic integrity** Boston College values the academic integrity of its students and faculty. It is your responsibility to familiarize yourself with the university's policy on academic integrity: www.bc.edu/integrity. If you have any questions, always consult your professor. Violations of academic integrity will be reported to your class dean and judged by the academic integrity committee in your school. If you are found responsible for violating the policy, penalties may include a failing grade as well as possible probation, suspension, or expulsion, depending on the seriousness and circumstances of the violation.

For more information, please make sure to read:
http://www.bc.edu/offices/stserv/academic/integrity.html