UIUC
Dept of Economics

Spring 2024
Econ 590 - MSPE
386 Armory, TR 9:30-10:50 (M4)

Applied Machine Learning in Economics

Instructor: Prof. Yannis Bilias (bilias@illinois.edu)

It's tough to make predictions, especially about the future. -Yogi Bera

Course Description

Machine learning, originally a development in computer science, combined with ideas of statistical analysis, offers the basis for a set of tools for modeling and understanding complex datasets.

This introductory course gives an overview of different concepts, techniques, and algorithms in machine learning with a view towards applications in economics. We begin with topics such as *Regression*, *Classification*, *Model Selection* and we move to methods of improved predictive accuracy like *Reguralized Regression*. Some more recent topics such as *Decision Trees*, *Boosting*, *Support Vector Machines*, and *Neural Networks* will be covered as time permits.

The course will be delivered from the vantage point of user and provide the student with skills of implementation of the basic machine learning methods in economic problems.

Prerequisite

Knowledge in business statistics at an intermediate level will be sufficient for the smooth attendance of the course.

Required Text

The main text of the course is:

- **(ISLR)** *An Introduction to Statistical Learning, (with Applications in R)*, 2nd ed, by James, Witten, Hastie, Tibshirani.

The text is freely available at: https://www.statlearning.com/

A version of the text with applications in Python is available:

- **(ISLP)** *An Introduction to Statistical Learning, (with Applications in Python)*, by James, Witten, Hastie, Tibshirani, Taylor.

- (Optional) For a deeper treatment on the fundamentals of statistical machine learning you can consult:

The Elements of Statistical Learning, by Hastie, Tibshirani, Friedman.

It is available at https://hastie.su.domains/ElemStatLearn/index.html

Computing
The course will make use of the *R programming language*. Previous experience with R is not required. For a detailed introduction to R language you can consult:

- *An Introduction to R*, by Venables, Smith and the R Core Team.

The text is available at: `https://cran.r-project.org/doc/manuals/R-intro.pdf`

Topics
A precise description of the material we cover in class and its exact reference to the textbook will be offered as we proceed. A tentantive list of the topics that will be covered is:

Week 1, Jan 16: Introduction to course, introduction to R (ISLR, chapter 1)

Week 2, Jan 23: Introduction to Statistical Learning (ISLR, chapter 2)

Week 3, Jan 30: Linear Regression (ISLR, chapter 3)

Week 4, Feb 6: Classification (ISLR, chapter 4)

Week 5, Feb 13: Resampling Methods (ISLR, chapter 5)

Week 6, Feb 20: Linear Model Selection, Regularization, PCA (ISLR, chapter 6),
Review - Prepare for the Midterm Exam.

Week 7, Feb 27: **Review** - Prepare for the Midterm Exam.
MIDTERM EXAM (Thursday February 29)

Week 8, Mar 5: Linear Model Selection, Regularization, PCA (ISLR, chapter 6)

Week 9, Mar 12: SPRING BREAK

Week 10, Mar 19: Tree-based Methods (ISLR, chapter 8)
Class Project proposal is due March 24.

Week 11, Mar 26: Tree-based Methods (ISLR, chapter 8)

Week 12, Apr 2: Support Vector Machines (SVM) (ISLR, chapter 9)

Week 13, Apr 9: Support Vector Machines (SVM) (ISLR, chapter 9)

Week 14, Apr 16: Unsupervised Learning with Clustering (ISLR, chapter 12)

Week 15, Apr 23: Deep Learning (ISLR, chapter 10) - if time permits.
Review - Prepare for the Final Exam.
Class Project is due April 28(?).

Week 16: **Review** - Prepare for the Final Exam.

Lectures and R Labs
Lectures explain theories and methodologies of machine learning methods. During the lectures, students will be exposed on the practical aspects of machine learning algorithms using the R language.

Exams and Assignments
There will be one midterm exam and one final exam on theories, concepts, and applied aspects discussed during the lectures. Students are also expected to submit assignments and a class project on which they apply the practiced methods to a real data-based problem. The answers to assignments should be submitted on *canvas* class platform **using Rmarkdown**.

Late Submission Policy
Assignments and final projects should be submitted by the deadlines. Late submissions will not be accepted unless there is a well-documented and verified reason.

Class Project
For the class project you should work as a *group no larger than 4 individuals*. **You are strongly encouraged to form the research groups of size 4 as early as possible in the semester.**

Each member of the group will receive the same points equal to the points given to the whole project report. The class project should involve the analysis of a relatively complicated data set (large number of observations, more than 5-6 variables) with the goal of investigating an interesting question using methods developed in class lectures.

Class Project Proposal
By March 24 you are asked to submit a one-page proposal for the class project. The proposal should include the problem you will research on, some preliminary descriptive analysis of the dataset you will use, and the methods you intend to utilize for analysis.

Office Hours and TA
Instructor's office hours (101A DKH): 3:30-4:30 Tues Thurs, or by request.

TA: Julian Oolman (jjpwade2@illinois.edu)
Office hours: Room 15 (Basement) DKH, 3:30-5:00 Thursday.

Grading Policy
The final grade of the course will be based on:
- 20% Assignments
- 20% Midterm Exam (February 29)
- 15% Class Project Proposals (Due March 24)
- 20% Class Project (Due April 28(?))
- 25% Final Exam