YILDIZ TECHNICAL UNIVERSITY
DEPARTMENT of ECONOMICS

MACHINE LEARNING IN ECONOMICS

(Thursday 10.00-12.50)

Syllabus

Instructor: Prof. Dr. Hüseyin Taştan
Email: tastan@yildiz.edu.tr
Classnotes: https://github.com/htastan/Machine-Learning-in-Economics
Office: Davutpaşa Campus, IIBF/G2-205
Office Hours: Tuesdays 14:00-16:30 or by appointment (please send an email)
Course assistant: tba

SCOPE and PURPOSE

The purpose of this course is to teach basic machine learning algorithms and methods relevant for empirical economists. The availability of large scale data sets has led to the development of new methods that are similar to those utilized in econometrics but also distinct in some aspects. This course, therefore, will emphasize the use of these new algorithms by focusing on statistical learning in economics and business. Students will learn the basic concepts, methods, and algorithms used in machine learning and develop skills to apply them in practice.

Econometrics software: We will use R in class and in lab sessions. R is an open-source software for statistical computing and graphics which is widely used by statisticians, researchers, data scientists and econometricians as well as industry professionals. The latest version of R can be downloaded from:
https://www.r-project.org/
And R-studio may be used as an integrated development environment for R:
https://www.rstudio.com/products/RStudio/

This class is supported by www.datacamp.com

PREREQUISITES

- You need to pass Statistics for Economists II (or equivalent). Prior knowledge of econometrics is not necessary but may be beneficial.

TEXTBOOK

- James, Gareth, D. Witten, T. Hastie, R. Tibshirani (2017), *An Introduction to Statistical Learning with Applications in R*, 8th ed., Springer.
  7th edition of this book is available freely in electronic format at the following address:
  https://www.statlearning.com/

- Alpaydın, Ethem (2016), *Introduction to Machine Learning*, 3rd ed., MIT Press, Cambridge MA.
  Here is the Turkish translation of this book:
- Alpaydın, Ethem (2018), *Yapay Öğrenme*, 4. Baskı (Ethem Alpaydın, Introduction to Machine Learning, 2. baskıdan çeviri), Boğaziçi Üniversitesi Yayınevi, İstanbul.

EVALUATION

Class attendance and participation: 10%, Homeworks and assignments: 30% , Midterm: 20%
Project and presentation: 40%

CLASS SCHEDULE

Week	Topics (Text: James et al.)	Preparation
1 (March 11)	Introduction to machine learning in economics, Concepts and tools in statistical learning theory, Lab: Introduction to R and RStudio	Ch. 1
2 (March 18)	Supervised and unsupervised learning, estimating prediction accuracy, bias-variance trade-off Lab: Introduction to R programming, handling data in base R	Ch.1, Ch.2, Ch. 5
3 (March 25)	Tidy approach to data analysis in R, introduction to the R Tidyverse	Classnotes
4 (April 1)	Introduction supervised learning methods Linear regression	Ch. 3
5 (April 8)	Classification problems, logistic regression, PCA, discriminant analysis	Ch. 4
6 (April 15)	Validation and Cross-Validation, bias-variance trade-off, Bootstrap, data-dependent information criteria	Ch. 5
7 (April 22)	Model Selection and Regularization: shrinkage methods, LASSO, Ridge regression	Ch. 6
8 (April 29)	Midterm	
9 (May 6)	Nonlinear extensions: polynomial regression, splines, local regression, General Additive Model	Ch. 7
10 (May 13)	No class	
11 (May 20)	Tree-based methods, decision trees, regression trees, random forests, bagging, boosting	Ch.8
12 (May 27)	Unsupervised Learning I: PCA, K-means clustering	Ch. 10
13 (June 3)	Unsupervised Learning II: Hierarchical clustering	Ch. 10
14 (June 10)	Projects due + presentations	