

**Boston College, Department of Economics**  
**ECON3389 Machine Learning for Economics**  
**Spring 2022**

Instructor: Haydar Evren, [haydar.evren@bc.edu](mailto:haydar.evren@bc.edu)  
Class days/times: Tue/Thu, 4.30pm-5:45pm  
Class location: O'Neill Library 257  
Office Hours: Mon/Wed/Fri, 3pm-5pm (Zoom only)

### **Course Overview**

One of the defining features of the world around us today is the ever-increasing amounts of data that describe our daily lives. This "big data" phenomenon, as it became known, has led to the development of new methods, called "machine learning," that allow high-dimensional statistical analysis in ways that were either impossible or infeasible for classical statistical methods, such as regression analysis.

This course aims to provide students with an introduction to modern data-driven learning in a framework that makes it applicable to causal economic analysis. While we will cover the necessary theoretical foundations, the emphasis will be placed on application and learning how and when to use these methods and when and how these methods can fail.

The coursework will include:

- Homework assignments with simulated and real-world data.
- Weekly online discussions on real-life data analysis issues.
- A group project in the form of a case study.

We will use R as our primary data analysis software, with a significant amount of class time devoted to learning how to code various analytical models efficiently. Prior coding experience is welcomed but not required — everything you need to know about R will be taught from scratch.

**Important:** This syllabus provides only the basic information about the course. The details on course structure, assignments, policies, and grading are provided in the introductory lecture. All students must familiarize themselves with both the syllabus and the introductory class.

### **Prerequisites**

For students majoring in Economics, Economic Statistics (ECON1151) is a required prerequisite, and Econometric Methods (ECON2228) is a recommended prerequisite. Students coming from different departments should have similar command of statistical methods. Solid knowledge of differential calculus at the level of MATH1102 (the "preferred" co-requisite for ECON 2228) is highly recommended. Prior knowledge of programming (Stata, R, Python, Matlab) will be helpful but is not a prerequisite; everything students need to know about coding in R will be taught from scratch in the course.

## **Literature**

Main textbook (required):

1. *An Introduction to Statistical Learning, with Applications in R*. James, G., Witten, D., Hastie, T., and Tibshirani, R (2013), available for free at: <https://www.statlearning.com/>

Additional introductory references (recommended):

2. *Machine learning with R Cookbook*. Chiu, Yu-Wei. (2015).
3. *Data Mining and Business Analytics with R*. Ledolter, J. (2013)
4. *A Gentle Introduction to Effective Computing in Quantitative Research*. Paarsch, H.J., Golyaev, K. (2016)
5. *Fundamentals of Machine Learning for Predictive Data Analytics*. Kelleher, J, Mac Namee, B. and D'Arcy, A. (2015)
6. *Learning from Data: A Short Course*. Abu-Mostafa, Y. , Magdon-Ismael, M. and Lin,

H. (2012) Additional advanced references:

7. *The elements of statistical learning: data mining, inference and prediction*. Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2009), available for free at <https://web.stanford.edu/~hastie/ElemStatLearn/>
8. *Statistical Learning with Sparsity: The Lasso and Generalizations*. Hastie, T., Tibshirani, R., Wainwright, M. (2015)
9. *Deep Learning*. Goodfellow, I., Bengio, Y., Courville, A., Bach, F. (2016)
10. *Machine learning: a probabilistic perspective*. Murphy, K. P. (2012).
11. *Pattern recognition and machine learning*. Bishop, C. M. (2006).
12. *Introduction to high-dimensional statistics*. Giraud, C. (2014).
13. *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Bühlmann, P. and Geer van de, S. (2011)

## **Required Software**

The primary software environment is the R statistical programming language, which can be downloaded for free from <http://www.r-project.org>. RStudio is the recommended interface for the R statistical programming language software also be downloaded for free at <http://www.rstudio.org>.

## **Canvas**

Canvas is the Learning Management System (LMS) at Boston College, designed to help faculty and students share ideas, collaborate on assignments, discuss course readings and materials, submit assignments, and much more - all online. As a Boston College student, you should familiarize yourself with this important tool. For more information and training resources for using Canvas, click [here](#).

All class materials will be available online via Canvas.

## Course components

Your course grade will be determined using the following components:

Homework	60%
Canvas discussions	10%
Case study	10%
Course project	20%

There are no letter grades or curving per each individual part of the grade. The overall course score is calculated as a weighted sum of all components, and then translated from 0-100 scale into a letter grade using a curved distribution. All students can access final grades through Agora after the grading deadline each semester. Transcripts are available through the [Office of Student Services](#).

## Deadlines, Late Work and Make Up Policy

All deadlines are strictly enforced. Late work is not accepted and no credit will be earned on late work unless the student has arranged an extension ahead of time with me (and that is quite possible, I am flexible with everyone's challenging circumstance and time constraints), with rare exceptions based on individual circumstances (e.g. inability to communicate with me ahead of time because of an emergency).

Please note that hardware ("WiFi was down") and software ("Canvas was glitching") issues are not considered valid excuses for late submissions.

## Course outline and schedule

Module	Topics	Assignments
Module 1	Introduction to R and RStudio	Discussion 01 Homework 01 Homework 02
Module 2	Statistical Learning	Discussion 02
Module 3	Linear Regression I	Discussion 03 Homework 03
Module 4	Linear Regression II	Discussion 04 Homework 04 Case study report (draft version)
Module 5	Classification	Discussion 05 Homework 05 Case Study Presentation Case Study Report
Module 6	Resampling methods	Discussion 06 Homework 06 Project report (draft version)
Module 7	Feature selection	Homework 07 Discussion 07
Module 8	Decision trees	Homework 08 Project Presentation Project Report

## **Accommodation and Accessibility**

Boston College is committed to providing accommodations to students, faculty, staff, and visitors with disabilities. Specific documentation from the appropriate office is required for students seeking accommodation in this course. Advanced notice and formal registration with the appropriate office are required to facilitate this process. There are two separate offices at BC that coordinate services for students with disabilities:

- [The Connors Family Learning Center \(CFLC\)](#) coordinates services for students with LD and ADHD.
- [The Disabilities Services Office \(DSO\)](#) coordinates services for all other disabilities.

Find out more about BC's commitment to accessibility at [www.bc.edu/sites/accessibility](http://www.bc.edu/sites/accessibility).

## **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. 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.

See [this link](#) for a full discussion of the university's policies and procedures regarding academic integrity.

## **Health Integrity Policy**

Particularly during this time of the COVID-19 pandemic, we must take even greater measures to care for ourselves, for each other and for our community. Therefore, all students are asked to care for themselves by monitoring their health and washing their hands thoroughly before class. At this time and subject to change, fully vaccinated faculty, staff, and students are not required to wear masks on campus, but should feel free to do so if they prefer, including in offices and classrooms. Those with approved University COVID-19 vaccination exemptions must wear masks outdoors and while in all campus buildings. Please continue to monitor the [BC Forward](#) site for updates for the BC community.