

# Computational Machine Learning I

Fall term 3 ECTS

Mandatory Course

Prof. Joan Verdú

## Prerequisites to Enroll

Although not mandatory, some knowledge of Python, Jupyter notebooks, and algebra is recommended.

## Overview and Objectives

This is an intensive 20-hour course based on a hands-on approach using Jupyter notebooks, all material is motivated by specific information retrieval and data analysis questions and each thematic unit concludes with a small project. The course introduces basic programming and analytics techniques.

The course will be delivered by Joan Verdú, Head of Consulting and Knowledge Transfer at the BGSE Data Science Center, in collaboration with Data Scientists affiliated to the Data Science Center.

## Course Outline

The course provides training in information retrieval, data analysis and machine learning with Python and R and evolves along the thematic following units:

1. Data exploration with Python

Keywords: pandas, database management, seaborn

2. Scraping data

Keywords: scraping, API's, cloud storage

3. Introduction to Data visualization

Keywords: pyplot, seaborn, plotly

4. Supervised learning

Keywords: sklearn, linear models, cross validation, regularisation, lasso, Auto MachineLearning

5. Unsupervised learning

Keywords: clustering, factors, independent component analysis and matrix factorization

# Computational Machine Learning I

Fall term 3 ECTS

Mandatory Course

Prof. Joan Verdú

## 6. Data exploration with R

Keywords: data analysis, data visualization, supervised and unsupervised models

## 7. Supervised and unsupervised learning with R

Keywords: linear models, clustering, cross-validation

### Required Activities

Class participation is compulsory. During classes there will be little projects that students are expected to perform while in class. There will also be larger projects that students are expected to solve in their own time.

### Evaluation

Attendance at classes, and submission of homework.

2 or 3 projects in total

- The in-class projects will be individual
- The home projects will be in groups of 2
- Will be given at the last class of each part
- Students will be given 2 weeks to submit their project.

The grade will be the average of these projects, plus eventually some extra tasks to be done in class.

### Competences

- Construct a global vision of the situation of the problem based on knowledge of the synergies between advanced statistical methods, computing and business analysis to generate added value.
- Modeling and predicting high-dimensional data with advanced statistical methods in the field of data science in order to improve strategic decision making.
- Apply the knowledge of programming languages, computer programs and advanced services in the Cloud to solve the problems that are presented to the data scientist.
- Solve the real problems that arise in the fields of study through the accurate analysis of the data.

# Computational Machine Learning I

Fall term 3 ECTS

Mandatory Course

Prof. Joan Verdú

- Visualize and interact with high-dimensional data in order to contextualize the information and facilitate subsequent decision-making.
- Communicate with conviction in English the results and implications of the required analytical study using a language related to the receiver.
- Work in a heterogeneous team of researchers in the field of the economic analyst using specific group techniques.
- Own and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context.
- That students know how to apply the acquired knowledge and their ability to solve problems in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study.
- That the students be able to integrate knowledge and face the complexity of making judgments based on information that, being incomplete or limited, include reflections on the social and ethical responsibilities linked to the application of their knowledge and judgments.
- That the students know to communicate their conclusions and the knowledge and last reasons that sustain them to specialized and non-specialized publics in a clear and unambiguous way.
- That students have the learning skills that allow them to continue studying in a way that will be largely self-directed or autonomous.

## Learning Outcomes

- Elaborate and estimate probabilistic prediction models based on certain data.
- Predict random processes.
- Apply supervised and semi-supervised learning algorithms.
- Apply search algorithms and estimation methodologies in networks through observation of data.
- Apply mathematical and computational analysis of social, business and economic networks knowing the theory and optimization algorithms.
- Predicting information needs based on decisions that must be made.
- Apply mathematical theory and statistics on data sets from disparate disciplines.

# Computational Machine Learning I

Fall term 3 ECTS

Mandatory Course

Prof. Joan Verdú

## Materials