

Applied Machine learning with Python

Lecturer: Andrea Giussani

Course language

English

Course description and objectives

The purpose of this course is to give students a solid introduction to modern applied Machine Learning (ML) methods and pipelines that are available for practitioners in the field of machine learning and statistical learning. In particular, the student will be guided through the construction of an end-to-end pipeline using both linear and nonlinear methods. In particular, each method will be briefly described and accompanied by hands-on, practical snippets written in Python. At the end of this course, each participant will be able to independently apply modern machine learning techniques to elementary problems related to Economics and Social Sciences.

Audience

The course is aimed for people involved in applying ML methods in the industry, especially for those interested in working in data analytics, or for those who want to extend their knowledge in modeling and statistical inference.

Prerequisites

The reader is expected to have a solid knowledge of Probability and Statistical Inference, and great familiarity with the Python language (or any other high-level, programming language, such as R or Julia). All students are strongly encouraged to bring their own laptop.

Duration

14 hours

Teaching mode

It will be possible to join the course in presence and/or in distance, by connecting remotely and following the streaming of the lesson held in the classroom.

Calendar

Lecture	Date	Time	Room	Lesson in person with groups by student ID number
1	Tue 16/02/2021	18.40 – 20.10	N29	Odd
2	Fri 19/02/2021	15.00 – 16.30	N29	Odd
3	Fri 19/02/2021	16.50 – 18.20	N29	Odd
4	Tue 23/02/2021	18.40 – 20.10	N29	Even
5	Fri 26/02/2021	15.00 – 16.30	N29	Even
6	Fri 26/02/2021	16.50 – 18.20	N29	Even
7	Tue 02/03/2021	18.40 – 20.10	N29	Odd

Syllabus of the course

Lecture	Topics	Book reference
1	Introduction to Machine Learning <ul style="list-style-type: none"> - Objectives of the course - Introduction to Machine Learning (ML) - The Python ML toolkit: a gentle introduction 	
2	Modern Machine Learning Pipeline: Preprocessing <ul style="list-style-type: none"> - Feature Distribution and Scaling - Normalization 	Ch. 1
3	Modern Machine Learning Pipeline: Preprocessing (2) <ul style="list-style-type: none"> - Imputation - Dealing with Categorical Variables - Model Selection 	Ch. 1
4	Modern Machine Learning Pipeline: Modeling <ul style="list-style-type: none"> - A simple ML Pipeline - Overfitting and Underfitting - Model Selection 	Ch. 2
5	Modern Machine Learning Pipeline: Modeling (2) <ul style="list-style-type: none"> - Linear vs Logistic Regression - Shrinkage Methods 	Ch. 2

6	Modern Machine Learning Pipeline: Modeling (3)	Ch. 3
	- Nonlinear Models: why do we need them	
	- Gradient Boosting	
	- Random Forest	

7	Final Exam and Discussion	
---	----------------------------------	--

Software used

Python 3.7 (or greater)

Suggested bibliography

- Giussani, A. – *Modern Applied Machine Learning with Python* – BUP (2020)
- Hastie, Tibshirani, Friedman - *The elements of statistical learning* – Springer (2009)

Available seats

This activity is limited to **60** participants and reserved to **students of the Master of Science Programs**. Registrations cannot be carried out once this number has been reached or after closing of the registration period.