



UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI ECONOMIA,
MANAGEMENT E METODI QUANTITATIVI



Centre of
Excellence in
Economics and
Data Science



SUMMER SCHOOL IN MACHINE LEARNING

AN INTRODUCTION TO MACHINE LEARNING USING STATA (ONLINE)

Date: July, 6 – 7 – 9 – 10 – 13, 2020

UNIVERSITY OF MILAN

*Department of Economics, Management and Quantitative Methods
Centre of Excellence in Economics and Data Science*

Instructor

Giovanni Cerulli

*Research Institute on Sustainable Economic Growth,
National Research Council of Italy, Unit of Rome*

giovanni.cerulli@ircres.cnr.it

Scientific Committee

Tommaso Frattini (Chair)

Emanuele Bacchiocchi

Massimiliano Bratti

Fabrizio Iacone

Giancarlo Manzi

Andreea Piri

Francesco Rentocchini

Silvia Salini

Organizing Committee

Anna Basoni (CEEDS)

Stefania Scuderi (DEMM)

Tutors

Andreea Piri (DEMM)

Veronica Rattini (DEMM)

Giuseppe Gerardi (CEEDS)

Course Content and Schedule

July, 6th Session 1

9.30 - 11.30	Fundamentals of Machine Learning
11.30 - 12.00	Break with questions preparation and collection
12.00 - 12.30	Replies to questions

July, 6th Session 2

14.30 - 15.30	Resampling and validation methods
15.30 - 16.00	Break with questions preparation and collection
16.00 - 16.30	Replies to questions

July, 7th Session 3

9.30 - 11.30	Model Selection and regularization
12:00 - 13:00	Computer lab
15.00 - 15.30	Replies to questions

July, 9th Session 4

9.30 - 11.30	Discriminant analysis and nearest-neighbor classification
12:00 - 13:00	Computer lab
15.00 - 15.30	Replies to questions

July, 10th Session 5

9.30 - 11.30	Tree-based methods
12:00 - 13:00	Computer lab
15.00 - 15.30	Replies to questions

July, 13th Session 6

9.30 - 11.30	Neural networks
12:00 - 13:00	Computer lab
15.00 - 15.30	Replies to questions

Additional Information

Software Requirements

Practical sessions will require the use of several software packages.

- Stata with the packages *lassopack*, *mlp2*, *subset*, *srtree*, *sctree*, *crossfold* (minimum software requirement is Stata version 14).
- R and RStudio (last available versions) with the packages *MASS*, *ISLR*, *foreign*, *plyr*, *tree*, *randomForest*, *gbm*, *leaps*.
- Python (version 3.7) with the scikit-learn package. The Anaconda distribution of Python is appropriate.

All sessions will be delivered through Zoom. Presentations and other course materials (including video recordings) will be shared through Dropbox links (available until August, 13th).

Preliminary Reading

- Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani (2013), *An Introduction to Statistical Learning with Applications in R*, Springer, New York, 2013.
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman (2008), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, second edition, Springer.
- Cerulli, G. (2020), “A Super-Learning Machine for predicting economic outcomes”, MPRA Paper 99111, University Library of Munich, Germany, 2020.

An in-depth introduction to machine learning (15 hours of videos) based on the textbook “An Introduction to Statistical Learning with Applications in R (ISLR)” by Trevor Hastie and Rob Tibshirani, is available here:

<https://www.r-bloggers.com/in-depth-introduction-to-machine-learning-in-15-hours-of-expert-videos/>

Administrative Issues

Three tutors were allocated to facilitate student learning. They will be answering students’ questions and will liaise with the course instructor in solving student issues.

The certificate of participation will be sent upon successful completion of the course. Minimum attendance requirements: 75% of the scheduled class meetings, daily attendance to the lectures and participation in the “Q&A” sessions. There are no assignments, yet students must partake in group activities that are scheduled to take place during computer lab sessions.

Other Arrangements

Interaction during class

The tutors – Giuseppe Gerardi, Veronica Rattini and Andreea Piriu – will participate to all sessions and will provide help with any technical issues that prevent students from moving on.

Students should mute their microphone during sessions. Questions can be submitted in the Zoom chat or through the “Raise your hand” option.

Q&A sessions

Every session is followed by a dedicated “Q&A” session. All questions asked in Zoom will be collected by the tutors and addressed during the “Q&A” sessions. Irrespective of having submitted a question in Zoom or not, all students are strongly encouraged to ask questions in these time slots.

Further information can be obtained by visiting:

<https://ceeds.unimi.it/summer-school-an-introduction-to-machine-learning-using-stata/>