

Machine Learning (F8203B040)

University of Milano Bicocca

General Information

Teachers (email: name.surname@unimib.it):

- Mirko Cesarini
- Stefano Peluso

Course Website <https://elearning.unimib.it/course/view.php?id=38389>

Office Hours by appointment (please, send an email to the teacher)

Accademic Year: 2021/22. **Credits:** 6CFU. **Language:** English

General Information	1
Content	2
Communications	2
Classes	2
Lecture Notes, Software, Papers, ...	2
Reference Books	3
Examination and Grading	3
Project Guidelines	3

Schedule

Classes are over for the Academic Year 2021/22.

First lecture: ~~Tuesday, March 1st, 2022 08:30-11:30 AM~~ GET (Central European Time) Lab716
—U7 Building. Lectures will usually be on Tuesday and Thursday, as outlined in the schedule
below. Class will be in ~~Lab716 (U7 Building)~~ if not otherwise specified.

- ~~Tue 08:30-11:30~~
- ~~Thu 11:30-15:30~~

First Lesson

- ~~March 1st, 2022 Tue 08:30-11:30~~
- ~~March 3rd, 2022 Thu 11:30-15:30~~
- ~~March 8th, 2022 Tue 08:30-11:30~~
- ~~March 10th, 2022 Thu 11:30-15:30~~
- ~~March 15th, 2022 Tue 08:30-11:30~~

Content

- Statistical methods for machine learning
 - Supervised and unsupervised learning
 - Recall to regression analysis
 - Classification analysis
 - Cross validation and bootstrap
 - Model selection and regularization
 - Beyond linear models
 - Tree-based methods
 - Support vector machines
- Feature Engineering and Machine Learning Algorithms Tuning
 - Feature Engineering and Selection (Bag of Words, Embeddings, Tensors, ...)
 - Data Observability and Model existence issues
 - Hyperparameters optimization (Grid-Search, Random-Search, Advanced Search methodologies)
- Artificial Neural Networks and Deep Learning
 - Artificial Neural Networks (ANNs) and Feed Forward Neural Network introduction
 - Training Algorithm: Gradient Descent, Optimization Methodology
 - Deep learning and Artificial Neural Networks types (Fully Connected networks, Feed Forward networks, Convolutional networks, Recurrent networks, ...)
 - Industrial applications and open research issues

Communications

Communication will be held by email. The teacher might use the course mailing list. Students enrolled in the course website will be automatically added to the course mailing list.

Classes

Lessons will be taught in person. Although students can attend remotely by video streaming, in-person participation is strongly suggested. The streaming link will be available on the course website (link on the first page) about half an hour before the lesson start.

Lecture Notes, Software, Papers, ...

All the material will be available on the course website (link on the first page)

Reference Books

- Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (2013). An Introduction to Statistical Learning. Springer. Available at <http://www-bcf.usc.edu/~gareth/ISL/>
- C.M. Bishop (2006), Pattern Recognition and Machine Learning. Springer (New York)

Examination and Grading

Oral exam. Students must carry on a project, individually or in groups of maximum three people. The project topic is to be discussed and agreed in advance with one of the teachers. Data sources, codes, and a report are to be handed in. Further information during classes.

Project Guidelines

The final project will count as 100% of your grade. It can be done in **1, 2, or 3 people**.

It will be an analysis of a **dataset** of your choice, using the methodologies learned during the course, or one related topic not directly faced in class. The topic has to be agreed in advance with one of the professors.

You will have to present your analysis and the **presentation** will count in the exam evaluation. The presentation will last about

- 15 minutes + 5 minutes for Q&A (1 person team)
- 20 minutes + 5 minutes for Q&A (2 people team)
- 25 minutes + 5 minutes for Q&A (3 people team)

During the presentation, there will be **questions** on topics covered during the course.

Material to turn in

- a written report;
- the R/Python commands/scripts used for the analysis;
- the data set used;
- the presentation.

More details on the **project structure** (the next bullets are suggestions, not compulsory, dependent on the specific topic):

- problem description and research questions
- data description, descriptive statistics, graphs
- it is also possible to work with simulated data

- the methodology used to answer the research question and discussion of any underlying hypotheses
- conclusions and further developments

This Document Link

https://docs.google.com/document/d/1G4eBH47YaPRIRH0nK7OC3Cvvqr_7dgPj1sV2VIL2ucQ/edit?usp=sharing