

## **MACHINE LEARNING AND APPLIED MACHINE LEARNING**

### # Learning Objectives

The aim of this course is to study and apply the most relevant data mining and machine learning models and algorithms to real cases datasets. The attendants will be able to distinguish between descriptive or explorative (unsupervised) and predictive (supervised) methods to extract valuable patterns and models from (large) data.

Important machine learning methods will be introduced theoretically and applied on real datasets using Python libraries.

The combination of theory and practice offers students the background necessary to choose the right tool for the job at hand, the expertise to autonomously apply them to real data sets, and the competence to measure the quality of the obtained results.

### # Course Content

- 1) Introduction: data, data preprocessing, distance metrics
- 2) Classification algorithms: logistic regression, decision trees, k-nn
- 3) Clustering Algorithms: k-means, hierarchical clustering, DBSCAN
- 4) Introduction to Neural Networks (NN)
- 5) Model Selection: general principles and applications to NN hyperparameters selection

### # Course Methodology

Frontal teaching and Lab.

### # Readings

Tan, Steinbach, Kumar: Introduction to Data Mining. Addison Wesley

Ian Goodfellow, Yoshua Bengio and Aaron Courville: Deep Learning

Slides: available during the course

## # Course Evaluation

Students will be evaluated on a final written exam (50%) and a python project (50%).

The written exam will be formed by 4 open questions on topic covered during the frontal lectures.

In the project they will set up a quick machine learning case study based on some datasets and tasks chosen by the students and will have to answer few questions related to how the students approached the learning problem and about the course program.

## # About the Instructors

Roberto Esposito received the M. Sc. Degree in Computer Science from the University of Turin in 1999 and then the Ph. D. in Computer Science in 2003. He is Assistant Professor at the Department of Computer Science, University of Turin. His main research interests range in several subfields of Machine Learning including Graphical Models, Ensemble Learning, Support Vector Machines, and Neural Networks. He regularly serves as a reviewer for prestigious journals and conferences such as TKDE and IJCAI. He is member of the Editorial Board of the Machine Learning journal. He authored about 60 research products, including international peer-reviewed conferences and journal papers.

Mirko Polato is an Assistant Professor at the Department of Computer Science of the University of Turin. Italy. Mirko Polato received his MSc and his Ph.D. in Brain, Mind, and Computer Science from the University of Padova (Italy) in 2013 and 2018, respectively.

In 2017, Dr. Polato was a visiting Ph.D. student at the Delft University of Technology in the Multimedia Computing group under the supervision of Prof. Martha Larson.

From 2018 to 2021, he was a post-doctoral fellow at the University of Padova, working on two H2020 projects.

He was Guest Editor of a Special Issue on Federated Learning for Frontiers in Artificial Intelligence. He served as a Program Committee member of several international conferences and as a referee for several international journals.

He authored about 50 research products, including international peer-reviewed conferences and journal papers.