

## Machine Learning in High Energy Physics

### Obiettivi

Scopo del corso è fornire ai partecipanti le nozioni di base del Machine Learning e la sua applicazione nell'ambito della fisica delle particelle. Le lezioni teorico-pratico saranno tenute da due docenti esperti internazionali e integrate da seminari di altri ricercatori dove si mostreranno casi concreti di applicazioni di ML e Deep Learning. Durante il corso, una frazione significativa del tempo a disposizione sarà impiegata per esercizi in sessioni hands-on. Al termine del corso, i partecipanti acquisiranno competenze di base circa i fondamenti probabilistici e statistici della teoria, gli algoritmi utilizzati per la catalogazione e riconoscimento di patterns e la loro applicazione in casi di interesse per la fisica delle particelle.

**Il corso sarà tenuto in lingua inglese.**

### Destinatari

Ricercatori e tecnologi coinvolti nell'analisi dei dati degli esperimenti.

### Date

20-22 Maggio 2019. Inizio ore 11 del primo giorno e termine ore 16 del terzo giorno.

### N. partecipanti

30.

### Responsabile:

Marco Battaglieri, Sezione di Genova (battaglieri@ge.infn.it) 010.3536736  
Pasquale di Nezza, Laboratori Nazionali di Frascati (dinezza@Inf.infn.it) 06.94032284

### Segreteria organizzativa:

Elisa Carnevali (elisa.carnevali@ge.infn.it), Debora Ferraro (debora.ferraro@ge.infn.it), Laura Opisso (laura.opisso@ge.infn.it) – Sezione di Genova

### Docenti:

Daniele Bonacorsi – Università di Bologna  
Amir Farbin – University of Texas at Arlington  
Cristiano Fanelli – MIT  
Alessandro Ferrari – Argo Vision  
Derek Glazier – University of Glasgow  
Jean-Roche Vlimant – CERN

### Streaming:

Sì  No

### Sede:

Hotel Cenobio dei Dogi Via N. Cuneo, 34 - 16032 Camogli – Genova.

## PROGRAMMA PRELIMINARE

May Monday 20 - Basics

11.00 - 11.30 Introduction

11.30 - 13.00 Lecture I (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

13.00 - 14.30 Lunch break

14.30 - 16.00 Lecture II (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

16.00 - 16.30 Break

16:30 - 17:30 Hands-on (preparation)

17.30 - 18.15 Automating data-analysis for CLAS12 (D.Glazier – University of Glasgow) (30+15mn)

18.15 - 19.00 Computer vision and applications (A.Ferrari – Argo Vision)(II (30+15mn)

May Tuesday 21 - Advanced

09:00 - 10.30 Lecture I (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

10.30 - 11.00 Break

11.00 - 12.30 Lecture II (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

12.30 - 14.30 Lunch break

14.30 - 15.30 Hands on (project I)

16.00 - 16.30 Break

16:30 - 17:30 Hands-on (project II)

17.30 - 18.15 Graph Network for High Energy Physics (J.R. Vilmant CERN) (30+15mn)

18.15 - 19.00 A.I. “wild-caught” data (C.Fanelli – MIT) (30+15mn)

May Wednesday 22 - More advanced + examples

09:00 - 10.30 Lecture I (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

10.30 - 11.00 Break

11.00 - 12.30 Lecture II (D.Bonacorsi - INFN-BO / A.Farbin University of Texas Arlington)

12.30 - 14.30 Lunch break

14.30 - 16.00 Closeout

In the lectures, an overview of the wide adoption of machine learning (ML) and deep learning (DL) techniques in science and beyond will be given, as well as an overview to the main ML/DL tools and framework used world-wide. A review of most common algorithms will be presented, including different kinds of neural networks (NN), like CNN and RNN, and their current and future usage in (selected) INFN research areas. The hands-on parts aim at putting the concepts into practice through examples. Starting from getting familiar with the Python language and useful libraries for ML (e.g. scikit-learn), the hands-on sessions will dive into a series of exercises of increasing complexity, using more complex real-world datasets, as well as exploring more complete and powerful tools and frameworks, like Keras with Google Tensorflow as a backend. The exercise will be prepared in such a way to allow a practical training at the course but also their continuation at home after the short training ends. The implication in terms of computing hardware (e.g. CPU vs GPU vs FPGA vs TPU) and the hidden technical debt in ML systems will also be briefly covered, as well as on-premise vs cloud access to computing resources for ML/DL at scale.

