

## Big data e Machine Learning ONLINE

### Obiettivi

Il corso ha l'obiettivo di introdurre le principali tecniche utilizzate nell'ambito del Machine Learning per analizzare grosse moli di dati (i cosiddetti Big data). Si introdurrà il concetto di Big data e si studierà come gestirli e processarli in maniera efficiente. Verranno studiati i principali modelli di ML (classificazione, clusterizzazione, regressione) e deep learning (vari tipi di reti neurali), e come gestire efficientemente il training e l'inferenza dei modelli su infrastrutture eterogenee (multicore, GPUs, cloud). Verranno altresì introdotte moderne tecnologie di automazione nell'ambito del Cloud Computing di livello PaaS, atte a semplificare l'uso dei framework necessari per l'esecuzione degli algoritmi di ML/DL. I partecipanti apprenderanno le "best practices" per applicare modelli di ML ad analisi di vario genere. Sono previste lezioni teoriche ed esercitazioni pratiche.

### Destinatari

Principalmente a ricercatori e tecnologi, ma in linea generale a tutti coloro che siano interessati alle dinamiche trattate e in particolare al personale dei servizi Calcolo.

### Requisiti del partecipante

necessaria la conoscenza di base di Python

### Date

11-15 ottobre 2021

### Orario

10-12 e 14-16 per ogni giornata di corso

### N. partecipanti

massimo 15

### Responsabile:

Federica Legger

### Segreteria organizzativa:

Danila Bortot - danila.bortot@to.infn.it

### Docenti:

Federica Legger (INFN TO), Sara Vallero (INFN TO), Giacinto Donvito (INFN BA), Marica Antonacci (INFN BA), Domenico Diacono (INFN BA) Alfonso Monaco (INFN BA)

### Streaming:

Sì  No

### Sede:

Online – Piattaforma ZOOM



## PROGRAMMA

### **Lunedì 11 ottobre**

- 10:00 - 11:00 Introduction to big data science
- 11:00 - 12:00 Feature extraction
- 14:00 - 15:00 Smart workflows and distributed computing beyond the grid paradigm
- 15:00 - 16:00 Esercitazioni

### **Martedì 12 ottobre**

- 10:00 - 11:00 Introduction to Machine Learning: Supervised models
- 11:00 - 12:00 Introduction to Machine Learning: Unsupervised models
- 14:00 - 15:00 Distributed data beyond the grid paradigm
- 15:00 - 16:00 Esercitazioni

### **Mercoledì 13 ottobre**

- 10:00 - 12:00 Introduction to Deep Learning and neural networks
- 14:00 - 15:00 Distributed infrastructures beyond the grid paradigm
- 15:00 - 16:00 Esercitazioni

### **Giovedì 12 ottobre**

- 10:00 - 12:00 Deep Learning architectures: CNN, Autoencoders, RNN
- 14:00 - 16:00 Esercitazioni

### **Venerdì 15 ottobre**

- 10:00 - 11:00 Introduction to INFN Cloud
- 11:00 - 12:00 Introduction to INFN Cloud services
- 14:00 - 16:00 Esercitazioni

Al termine del corso i partecipanti realizzeranno una breve presentazione finale relativa ad alcuni aspetti trattati durante le esercitazioni.