



Intelligenza Artificiale e Supercalcolo

Come utilizzare tecniche di data analytics scalabili in azienda

3-4 marzo 2022

Torino

Ti stai chiedendo come risolvere i tuoi problemi di data analytics in tempi rapidi? Hai un dataset di dimensioni tali da non riuscire ad analizzarlo sulla tua workstation? In questo corso analizzeremo un dataset di grandi dimensioni sfruttando le moderne tecniche di Machine Learning e mostreremo come scalare un algoritmo, ovvero parallelizzarlo nei punti chiave (ad es. data pre-processing, cross validation) per sfruttare contemporaneamente le risorse di calcolo a disposizione (dal laptop al supercomputer).

Topics

1. Analisi dati e Machine Learning
2. Supercalcolo e supercomputers
3. Parallelizzazione di algoritmi
4. Supercalcolo per l'Intelligenza Artificiale

A chi è rivolto

Lingua: Italiano

Personale aziendale con background tecnico-scientifico, competenze base di Intelligenza Artificiale e conoscenza del linguaggio di programmazione Python, interessate a conoscere diverse tecniche di parallelizzazione degli algoritmi di Machine Learning, su workstation e su supercomputer.

Prerequisiti

(Competenze che devono avere gli studenti)

- i. Conoscenza di Python.
- ii. Familiarità con algoritmi di Machine Learning.
- iii. Capacità di eseguire codice su macchina Linux.
- iv. Conoscenze base di Machine Learning e Reti Neurali

Competenze in uscita

- i. Conoscenze di esempi di algoritmi di machine learning e gestione dati
 - ii. Conoscenze benefit HPC per ML e AI
 - iii. Conoscenze relative ai training e gli hyperparameter
 - iv. Conoscenze su come eseguire AI su HPC
 - v. Conoscenze teoriche e pratiche di training distribuito
-

DOCENTI

BHASKAR AGARWAL

CINECA

Ha conseguito il dottorato in Astrofisica Teorica al Max Plank Institute per la fisica extraterrestre di Monaco di Baviera (Germania). Dopo aver lavorato come ricercatore a YALE (USA) e Heidelberg (Germania), è diventato Data Scientist presso la start-up OmegaLambaTec. Lavora ora come Data Scientist presso il dipartimento HPC di Cineca, dove si occupa di algoritmi di Intelligenza Artificiale applicati all'Industria e a progetti di ricerca.

ERIC PASCOLO

CINECA

Ha conseguito il Master in High Performance Computing presso la SISSA di Trieste e ha poi lavorato come ricercatore HPC presso OGS nel campo dell'oceanografia. Successivamente ha iniziato a lavorare in CINECA, dove oggi ricopre il ruolo di consulente per le applicazioni industriali e project manager tecnico per i progetti di trasferimento tecnologico verso le industrie e in particolare le SME. Si è sempre occupato anche di formazione e divulgazione su tematiche legate al supercalcolo, sia per le imprese che per il grande pubblico. È inoltre responsabile del training del centro di competenza HPC nazionale EUROCC.

LAURA MORSELLI

ASSOCIAZIONE BIG DATA

Laureata in Astrofisica e Cosmologia all'Università di Bologna, ha conseguito un dottorato in Astrofisica alla Technische Universitat di Monaco di Baviera nel campo dell'evoluzione delle galassie. Lavora ora come HPC Specialist all'Associazione Big Data di Bologna.

TIMELINE

GIORNO 1

9:00	- Introduzione al Machine Learning e ai principali algoritmi per supervised e unsupervised learning.
11:00	Break
11:15	- Introduzione all'hands-on: supervised & unsupervised
13:00	Lunch
14:00	- Introduzione al parallelismo - Hands-on: // on workstation (data pre-processing)
16:00	Break
16:15	- Hands-on: // on workstation (hyperparameters tuning/cross validation)
18:00	End

GIORNO 2

9:00	- Interfacciarsi con un supercomputer: scheduler, jobscripts - Hands-on: porting dell'esercizio
11:00	Break

11:15	- Intro: Deep Learning & Neural Networks
13:00	Lunch
14:00	- Data Parallelism vs Model parallelism - Horovod
16:00	Break
16:15	- Hands on Horovod - Open Discussion
18:00	End