MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

Contribution ID: **253** Type: **not specified**

High-Performance Unsupervised Machine-Learning in Numerical Simulations and Satellite imaging

Thursday, 11 July 2024 17:10 (20 minutes)

Unsupervised machine-learning (UML) techniques play a crucial role in data analytics: clustering algorithms, in particular, are widely adopted in Astronomy. We ported a UML algorithm, the Advanced Density Peak (ADP) from the original Python code to an optimized C code and parallelized it with OpenMP, speeding it up by a factor of 40. We then applied ADP, which couples extreme sensitivity to small signal-to-noise and limited runtimes, to two astrophysical problems: (i) finding substructures in cosmological numerical simulations and (ii) de-blending images from the newly launched EUCLID satellite. We discuss these applications and review the preliminary results. Furthermore, the constant increase in the size of the datasets makes it already impossible to fit the data from cutting-edge experiments on a single node. That is why we are developing a new high-performance distributed-memory parallelization for ADP using MPI, which we also illustrate and discuss as of general interest in this field.

Presenter: TOMBA, Francesco

Session Classification: SPACE CoE