Responsabile Scientifico:

Prof. Alexandro Saro

co-responsabile: Dr. Matteo Costanzi

co-responsabile: Prof. Stefano Borgani

Titolo progetto di ricerca:

Cosmology with Machine Learning techniques:

Abstract:

The project will focus on the development and application of novel machine learning (ML) techniques to astronomical data for their cosmological exploitation, with focus on clusters of galaxies. Specifically, the PhD student will employ ML algorithms to classify cosmological structures and build catalogs of galaxy clusters from data, identify robust and unbiased proxies for the cluster mass, as well as to derive cosmological constraints from the statistical properties of the cluster population.

Upcoming wide field survey, such as Euclid, LSST-Rubin, eRosita, and CMB-S4, will provide us with a huge pile of data with the capability of shedding light on many fundamental questions in cosmology such as the nature of dark matter and dark energy. However, to unlock the full cosmological information encoded in these data sets will require an improved control over systematics and an extension of our current predictive models beyond summary statistics (such as number counts or correlation functions). ML algorithms represent a novel and powerful tool with the potential of overcoming many of these limitations, thanks to their capability of learning patterns and features from complex data sets. The development and application of ML-based algorithms to cosmological data, which are unique in terms of size and complexity, will lead the PdD student to develop skills in this strategical field which will have countless applications outside academia.

For cluster cosmology studies, which are currently limited by our capability of building catalogs with well defined selection functions and mass proxies, ML algorithms can help to improve several ingredients of the analysis pipeline : from the cluster detection to cosmological inference, passing through the definition of optimal mass proxies and modelling of systematic effects.

The student will work at the Department of Physics, also in close collaboration with INAF-Trieste, and have continuous interactions within an already well established group of experts. The candidate is expected to develop tools and expertise in coding innovative algorithms and analyzing both observational and simulated large datasets, including existing proprietary data (e.g. Dark Energy Survey, South Pole Telescope, and Euclid) and state of the art cosmological simulations developed within our group. The PhD student will join international collaborations (DES, Euclid, LSST) which will put him in contact with a stimulating and competitive research envirorment. The candidate is expected to participate at

least to one PhD school, present the outcomes of her/his work at national and international conferences and publish two/three high impact papers in peer-reviewed journals.

Foreseen timeline:

- In the first year the candidate will learn the theoretical and technical skills required for the development of the project; the student will start getting familiar with different ML architectures (e.g. random forest, neural networks) for parameter regression and classification, as well as with the astrophysical data needed to carry out the research project (e.g. photometric and mm cluster data).
- The second year will be devoted to the development and optimization of ML algorithms for the cluster detection and mass estimate from photometric and/or mm images. This testing phase will be mostly carried out on simulated data, assuming different levels of realism to ensure full control of the systematics.
- During the third year the student will capitalize his/her findings by applying the developed machinery to observational data (e.g. DES, Euclid, SPT) to produce well characterized cluster catalogs with robust mass proxies optimized for the cosmological analysis.