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

Contribution ID: **189** Type: **Oral Presentation**

Galaxy Morphological Classification via Unsupervised Machine Learning in the Big Data Era led by JWST, EUCLID, LSST and SKA

Tuesday, 9 July 2024 12:40 (20 minutes)

The morphological properties of galaxies are important tracers of the physical processes, e.g. minor/major mergers, gas accretion and tidal interactions, that have shaped their evolution. Forthcoming 'Big data'surveys (e.g. LSST/SKA), which will produce exabyte volumes of data will be the new 'normal'in this decade. These volumes will make morphological classification using traditional methods (e.g. direct visual inspection) impractical. Even semi-automated techniques, e.g. supervised machine learning with training sets built via visual inspection, may be difficult, because of the time-consuming nature of creating the training sets. However, unsupervised machine learning, does not require training sets, making it ideal for galaxy morphological classification for large surveys.

We present an unsupervised machine learning algorithm, that utilizes hierarchical clustering and growing neural gas networks to group together survey image patches with similar visual properties, followed by a clustering of objects (e.g. galaxies) that are reconstructed from these patches. We implement the algorithm on the Deep layer of the Hyper Suprime-Cam Subaru-Strategic-Program, to reduce a population of hundreds of thousands of galaxies to a small number (~100) of morphological clusters, which exhibit high purity. These clusters can then be rapidly benchmarked via visual inspection and classified by morphological type showing general morphologies (e.g. elliptical, spirals) but also peculiar/rare objects (e.g. mergers, clumpy discs). Using these morphological clusters, we successfully reproduce many known trends of galaxy properties (e.g. stellar-mass functions, rest-frame colours) as a function of morphological type. Furthermore, we find a rare type of elliptical galaxy in the dwarf regime ($10^8<M<10^{9.5}$ M_SUN) which exhibits the same colors and SFR activity as spirals but resembles elliptical shapes and Sersic profiles (blue elliptical) which we discuss in Lazar et al. (2023).

Due to its excellent ability to produce fine morphological classifications of even rare objects with minimal human intervention, in forthcoming papers we will use this algorithm on other future surveys as well (e.g. Euclid, JWST) which will give us the chance to explore in detail the morphological evolution of galaxies as a function of redshift, stellar mass and local environment across 80 per cent of cosmic time from an unprecedented statistical perspective.

Presenter: LAZAR, Ilin

Session Classification: Past and future multiwavelength all-sky surveys