

---

# Unsupervised Learning for Galaxy Cluster Membership Assignment

**Keywords** – Astrophysics - Cosmology - Machine Learning

---

The CosmoStat group at CEA Paris-Saclay invites applications for a 4-6 month internship to work with Dr. Samuel Farrens and Dr. Joana Frontera-Pons on machine learning and cosmology.

CEA Paris-Saclay is located 20km south of Paris, France, in the vicinity of various universities and other research centres. The CosmoStat group is a diverse and multi-disciplinary team of researchers working on various topics in cosmology. Our group is committed to diversity and equality, and encourage applications from women and underrepresented minorities. We support a flexible and family-friendly work environment.

## Context

Upcoming astrophysical surveys such as Euclid<sup>1</sup> aim to tighten cosmological constraints using properties of galaxy clusters such as their mass function and clustering strength. Optical cluster detection algorithms often provide unreliable or no information as to which clusters individual survey galaxies belong. This information significantly impacts the determination of cluster density profiles and velocity dispersion measurements. Current techniques for assigning galaxies to clusters generally involve defining arbitrary cuts in colour and redshift and may not take into account all of the available information.

Unsupervised machine learning techniques offer the flexibility to incorporate all available features in the data and to potentially identify better correlations between objects. Notably, clustering methods aim at creating groups or clusters so galaxies in the same cluster have similar properties, and galaxies in different groups are considerably distinct. Thereupon, clustering approaches strongly rely on the definition of a distance between two different objects. This project will focus on the development of new clustering techniques and on the selection of the galaxy features that properly define a cluster.

---

## Outline of project objectives

The internship will essentially be comprised of the following tasks and objectives:

1. Get familiarised with the current state of the art for galaxy cluster membership assignment.
  2. Get familiarised with unsupervised machine learning methods.
  3. Test this framework on simulated galaxy catalogues and compare with other methods.
  4. Apply the framework to real data and study the impact this has on cluster properties.
- 

## Candidate

The candidate should be a Master 2 (or equivalent) student with background in either physics/astrophysics or applied maths/signal processing/data science. Knowledge of machine learning methods would be a plus. Experience with Python is not required, but would be advantageous.

---

<sup>1</sup><https://www.euclid-ec.org/>

## Internship

The internship will take place in the CosmoStat laboratory, under the supervision of Samuel Farrens and Joana Frontera-Pons.

- *Deadline for applications:* December 12th, 2020.
  - *Contact:* Samuel Farrens ([samuel.farrens@cea.fr](mailto:samuel.farrens@cea.fr)), Joana Frontera-Pons ([joana.frontera-pons@cea.fr](mailto:joana.frontera-pons@cea.fr)).
  - *Duration:* 4-6 months.
-