UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

DEPARTAMENTO DE FÍSICA

Defensa Proyecto de Investigación para optar al grado de Licenciado en Astrofísica.

Gerardo Alonso Rodriguez Jorquera

Tema: ***Identifying Galaxy Cluster Members with Machine Learning out to Large Cluster-centric Distances***"

Comisión:

- Dra. Yara Jaffé, profesora guía
- Dra. Pía Amigo
- Dr. Rory Smith

Lunes 5 de enero a las 11:00 hrs.
Sala Conferencias Dr. Luciano Laroze, E300

Abstract:

Galaxy clusters are the largest gravitationally bound structures in the Universe and serve as powerful laboratories for studying galaxy evolution in dense environments. However, accurately identifying cluster members is a challenging task due to projection effects and the presence of interlopers, particularly in the cluster outskirts beyond the virialized core. This work addresses the membership classification problem out to large cluster-centric distances ($5\,r_{200}$) using a novel Machine Learning approach within the context of the CHANCES survey.

We utilize CHANCES mock catalogs constructed on the Uchuu simulation to train a Random Forest classifier, employing a sample of 42 massive clusters $\left(M_{200} > 10^{14}\,\mathrm{M}_{\odot}\right)$ at low redshift $z < 0.07$. To address the inherent class imbalance, we evaluate six model variants by combining three resampling strategies (Original Data, SMOTE and RandomUnderSampler), with and without hyperparameter optimization, training the algorithms on Projected Phase Space information $(v_{pec},\ r_{200})$ combined with local density (Σ_{10}) and intrinsic galaxy and halo information $(m_{r},\ M_{200})$.

Our best-performing model achieves an overall Completeness of 93% and a Purity of 66%, maintaining an optimal balance. We observe a radial dependence on performance, where purity naturally decreases towards the outskirts due to severe background contamination. The analysis reveals that the algorithm relies heavily on dynamical features, behaviimilarly to the traditional caustic technique and yielding competitive results with state-of-the-art approaches. Our findings demonstrate that Machine Learning is a robust tool for this task,ffectively recovering members at large radii, where the limitations are primarily physical—driven by projection effects—rather than algorithmic, establishing a reliable framework for future analyses of the CHANCES spectroscopic data.