Machine Learning Approach for the Spectral Classification of Dusty Stellar Sources in the Magellanic Clouds

Sepideh Ghaziasgar[1], Mahdi Abdollahi[1], Atefeh Javadi[1], Jacco Th. van Loon[2], Iain McDonald[3], Joana Oliveira[2], Amirhossein Masoudnezhad[4], Habib G. Khosroshahi[1], Bernard H. Foing[5], and Fatemeh Fazel[5]

[1]School of Astronomy, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[2]Lennard-Jones Laboratories, Keele University, ST5 5BG, UK
[3]Jodrell Bank Centre for Astrophysics, Alan Turing Building, University of Manchester, M13 9PL, UK
[4]Department of Physics, Sharif University of Technology, Tehran, Iran
[5]Leiden Observatory, Leiden University, 2300 RA Leiden, Netherlands

Abstract. We utilized supervised machine learning algorithms to classify dusty stellar point sources in the Magellanic Clouds, including young stellar objects (YSOs) and evolved stars. The classification accuracy of these sources was assessed using 12 multiwavelength features and 619 stellar objects' spectral labeled data derived from the Surveying the Agents of Galaxy Evolution (SAGE). We implemented various models to classify dusty stellar categories comprising YSOs, oxygen- and carbon-rich asymptotic giant branch stars (AGBs), red supergiants (RSGs), and post-AGB (PAGB). In addition, we created synthetic samples based on the minority class using the SMOTE method to resolve the imbalanced and small SAGE catalogs. As a result, the Probabilistic Random Forest (PRF) classifier performed the best with the highest accuracy among all the models applied before and after data augmentation, reaching **89**% based on the recall metric in categorizing dusty stellar sources.