EXTENDING GAIA XP SPECTRA BASED CLASSIFICATION OF HOT SUBDWARFS WITH MULTIPLE MACHINE-LEARNING APPROACHES AND AUXILIARY STELLAR DATA

Urtė Gedvilaitė[1], Erika Krikščiūnaitė[1], Aidas Medžiūnas[1], Kristupas Pribušauskas[1]

[1]Vilnius University, Faculty of Mathematics and Informatics, Institute of Applied Mathematics, Vilnius, Lithuania
kristupas.pribusauskas@mif.stud.vu.lt

This work is built on the Gaia XP spectra based machine-learning framework for hot subdwarf identification and classification introduced by Ambrosch et al. [1] Using this analysis as a reference point, the performance and complementarity of several additional machine-learning algorithms, including random forests, gradient-boosted trees, support-vector machines, and neural-network–based models, applied to Gaia XP spectral representations, are explored. Gaia XP spectra data remain the primary input, while the feature space is expanded to include selected astrometric and photometric parameters from Gaia as well as external information for stars of interest from complementary surveys.

Classification behaviour across algorithms is compared, robustness against known contaminants is assessed, and how alternative models and added data sources affect class boundaries is examined. Performance is evaluated using standard metrics and visualisation techniques, enabling direct comparison with the reference study. The results demonstrate how combining Gaia XP spectra with diverse machine-learning methods and auxiliary stellar data can refine classification confidence and uncover additional candidate objects for further investigation.

Keywords: hot subdwarfs, classification, machine learning, GAIA, spectra

[1] M. Ambrosch et al., “Detection of hot subdwarf binaries and sdB stars using machine learning methods and a large sample of Gaia XP spectra,” arXiv (Cornell University), Jan. 2026, [Online]. Available: http://arxiv.org/abs/2601.21727