

A Machine Learned Classifier for RR Lyrae in the VVV Survey

Felipe Elorrieta Lopez^{1,2}, Susana Eyheramendy^{1,2}, Andrés Jordán^{1,2} *et al.*

felorrieta@mat.puc.cl

¹ *Pontificia Universidad Católica de Chile.*

² *Millennium Institute of Astrophysics, Santiago, Chile*

Resumen

Variable stars of RR Lyrae type are a prime tool with which to obtain distances to old stellar populations in the Milky Way. One of the main aims of the Vista Variables in the Via Lactea (VVV) near-infrared survey is to use them to map the structure of the Galactic Bulge. Owing to the large number of expected sources, this requires an automated mechanism for selecting RR Lyrae, and particularly those of the more easily recognized type ab (i.e., fundamental-mode pulsators), from the variables expected in the VVV survey area. In this work we describe a supervised machine-learned classifier constructed for assigning a score to a Ks-band VVV light curve that indicates its likelihood of being ab-type RR Lyrae. We describe the key steps in the construction of the classifier, which were the choice of features, training set, selection of aperture, and family of classifiers. We find that the AdaBoost family of classifiers give consistently the best performance for our problem, and obtain a classifier based on the AdaBoost algorithm that achieves a harmonic mean between false positives and false negatives of $\approx 7\%$ for typical VVV light-curve sets. This performance is estimated using cross-validation and through the comparison to two independent datasets that were classified by human experts.

Keywords: stars: variables: RRLyrae - methods: data analysis - methods: statistical - techniques: photometric