RevMexAA (Serie de Conferencias), **49**, 97–97 (2017)

XV Latin American Regional IAU Meeting (Cartagena de Indias, Colombia, 3-7 October 2016)
Editors: Mario Armando Higuera Garzón & Santiago Vargas Domínguez

AN APPLICATION OF SUPERVISED LEARNING METHODS TO SEARCH FOR VARIABLE STARS IN A SELECTED FIELD OF THE VVV SURVEY

B. Rodríguez-Feliciano[1], A. García-Varela[1], M. F. Pérez-Ortiz[1], B. E. Sabogal[1], and D. Minniti[2]

We characterize properties of time series of variable stars in the B278 field of the VVV survey, using robust statistics. Using random forest and support vector machines classifiers we propose 47 candidates to RR Lyraæ, and 12 candidates to WU Ursæ Majoris eclipsing binaries.

Supervised learning methods have been used to search for variable stars, based on the properties of the light curves. These methods develop the process of classification using a previous classified training set. The algorithm we used, is built by choosing statistic properties of ell elements of the training sample. The supervised methods used were k-nearest neighbors, support vector machines, classification trees, and random forests.

We used near infrared data of a selected field in the Galactic Bulge, collected by the Vista Variables in the Via Lactea Survey. For the set of magnitudes in the K_s band of each star, we computed the robust statistics: median, median absolute deviation, octile skewness and tails weight. The statistics selected do not depend on a calculation of the period (obtained via Fourier analysis, analysis of variance or any other algorithm), so the learning methods are an alternative for the conventional methods. These were selected according to similar studies in our astro-statistics investigation group (Pérez-Ortiz et al. 2017).

To start we performed a usual search for different types of variable stars and propose 49 candidates for RR Lyraæ, 5 for Cepheids, 6 Algol binaries, 7 β Lyræ binaries (see examples of light curves in Fig. 1), and 25 WU Ursæ Majoris binaries. We used these stars as training sample to developed binary classifiers for the RR Lyræ and WU Ursæ classes. Using the most accurate supervised learning methods developed to detect variables stars (random forest and support vector machines), we propose 47 candidates to RR Lyræ, and 12 candidates to WU Ursæ Majoris. These classifiers had an accuracy of 80 % (see Fig. 2), estimated using cross validation between partitions of the training sample reevaluated in the classifier, after being obtained via R programing language using previous developed algorithms (Kuhn et al. 2016).

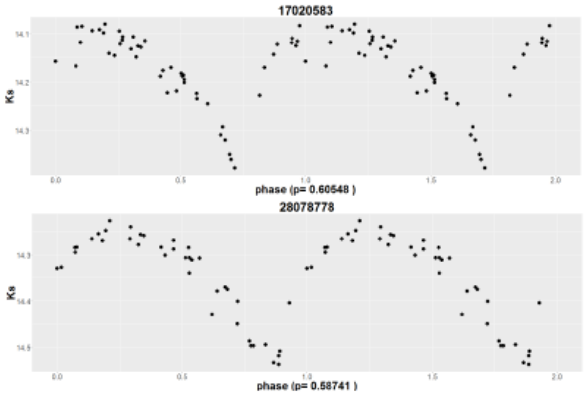

Fig. 1. Representative K_s -band light curves of RR Lyræ star candidates, in the training sample (top) and proposed by the classifiers (bottom). Their ID corresponds to assigned in the photometry of the template epoch (September 15 2011). Their period is indicated in days.

	Accuracy	Sensitivity	Specificity
Knn	0.7444444	0.04000	1.00000
Trees	0.7761111	0.4400	0.8971
Forests	0.8097222	0.4400	0.9412
SVM	0.8094444	0.32000	0.98529

Fig. 2. Accuracy, sensitivity and specificity of the binary classifiers obtained for the W Ursae majoris.

REFERENCES

M. Kuhn et al. The package caret: Classification and Regression Training, Online Package Manual. 2016.
Pérez-Ortíz et al. 2017, submitted to A&A.

[1]Universidad de los Andes, Departamento de Física, Cra. 1 No. 18A–10, Bloque Ip, A.A. 4976, Bogotá, Colombia (bs.rodriguez128, josegarc, mf.perez648, bsabogal@uniandes.edu.co).
[2]Universidad Andrés Bello, Departamento de Ciencias Físicas, Av. República 220, Santiago, Chile.