

Streaming classification of variable stars

Zorich, L., Pichara, K., Protopapas, P. (2020). Streaming classification of variable stars. *Monthly Notices of the Royal Astronomical Society*, 492(2), 2897–2909, <https://doi.org/10.1093/mnras/stz3426>

ABSTRACT

In the last years, automatic classification of variable stars has received substantial attention. Using machine learning techniques for this task has proven to be quite useful. Typically, machine learning classifiers used for this task require to have a fixed training set, and the training process is performed offline. Upcoming surveys such as the Large Synoptic Survey Telescope will generate new observations daily, where an automatic classification system able to create alerts online will be mandatory. A system with those characteristics must be able to update itself incrementally. Unfortunately, after training, most machine learning classifiers do not support the inclusion of new observations in light curves, they need to re-train from scratch. Naively re-training from scratch is not an option in streaming settings, mainly because of the expensive pre-processing routines required to obtain a vector representation of light curves (features) each time we include new observations. In this work, we propose a streaming probabilistic classification model; it uses a set of newly designed features that work incrementally. With this model, we can have a machine learning classifier that updates itself in real time with new observations. To test our approach, we simulate a streaming scenario with light curves from Convention, Rotation and planetary Transits (CoRoT), Orbital Gravitational Lensing Experiment (OGLE), and Massive Compact Halo Object (MACHO) catalogues. Results show that our model achieves high classification performance, staying an order of magnitude faster than traditional classification approaches.