Analysis of Kepler light curves using the wavelet transform to discriminate with machine learning the astrophysical nature of the eclipsing object.

Luis Guirado Fuentes, Roberto Baena Gallé - UNIVERSIDAD INTERNACIONAL DE LA RIOJA (UNIR)

XV REUNIÓN CIENTÍFICA SEA

INTRODUCTION

The Kepler mission has been the most successful so far in the search for and characterization of exoplanets using the transit technique. With this method, the intensity of light emitted by the star is measured at regular intervals to detect periodically recurring photometric reductions in the star, from which the presence of an eclipsing object can be inferred.

The wavelet transform has been used as an alternative to the Fourier transform in noise filtering in astronomical photometric data, as well as in the detection of exoplanet transits. We propose a new approach based on the use of the wavelet transform as a mathematical tool to establish statistical criteria for the characterization of the eclipsing object, in order to differentiate exoplanets from false positives, with the aim that the results obtained can be used to train a ML model to automatically analyze thousands of light curves Kepler and K2 missions.

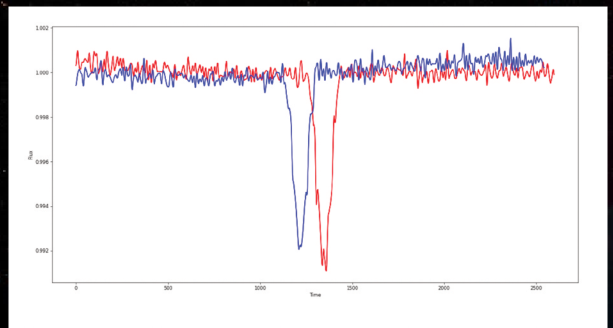

Figure 1: Representation of the folded light curve of the binary (false positive) KIC 10848459 system filtered by the wavelet transform to remove high frequencies. Odd transit represented in red and even transit in blue.

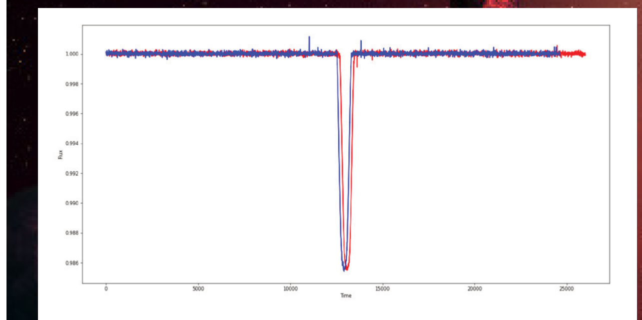

Figure 2: Representation of the folded light curve of the object KIC 11446443 filtered by the wavelet transform to remove the high frequencies, with the transits of the exoplanets Kepler-229 c. Odd transit in red and even transit in blue.

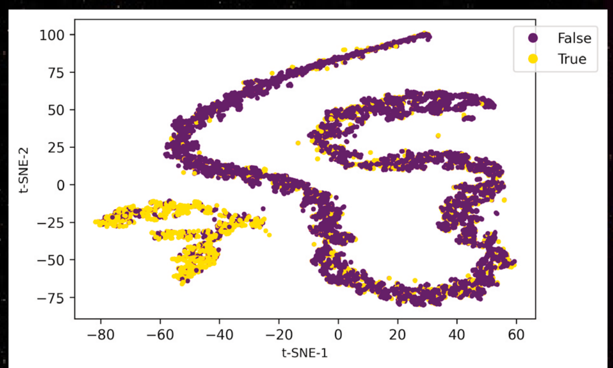

Figure 3: Visualization of the dimensionality of the clusters generated by the t-SNE method over the features dataset extracted with tsfresh. False positives in violet.

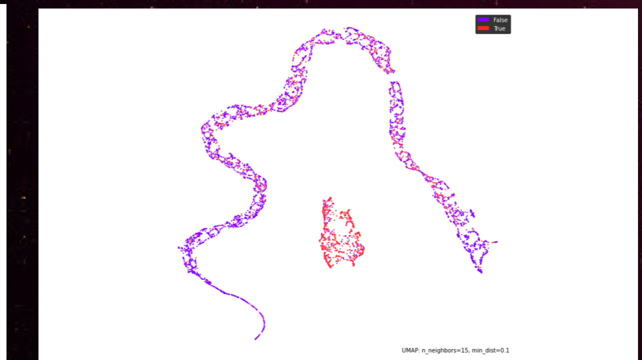

Figure 4: Visualization of the dimensionality of the clusters generated by UMAP method over the dataset of parameters extracted with tsfresh. False positives in blue.

METHODOLOGY

The methodology followed to process the Kepler Mission light curves is as follows:

- Computation of the phase-folded curve and extraction of the odd and even mask.
- 7-level discrete wavelet transform [1] "Symlet 5" applied to the odd and even data.
- Reconstruction of the signal by removing the high frequency components. Figures 1 and 2.
- Parameter extraction with tsfresh [2].
- LightGBM [3] processing and visualization of clusters through t-SNE[4] and UMAP[5]. Figures 3 and 4.

RESULTS/CONCLUSIONS

The results show that the implemented method allows to extract useful information to discriminate, in an automatic way, binary eclipses that are usually detected as exoplanets (false positives), with an accuracy of 81%. Further evaluation of more complex extracted features and enlargement of the data set through simulated curves is required.

The present work is a first approach to the identification of the nature of the eclipsing objects present in the light curves of missions such as Kepler by extracting features from these light curves processed and filtered through the wavelet transform.

REFERENCES

[1] Torrence, C. y Compo, G. P. (1998). A *Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society*, 79(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079%3C0061:apgtwa%3E2.0.co;2

[2] *tsfresh documentation.* (s. f.). https://tsfresh.readthedocs.io/

[3] *LightGBM 3.3.2 documentation.* (s. f.). https://lightgbm.readthedocs.io/en/v3.3.2/

[4] *t-SNE.* (s. f.). Laurens van der Maaten. https://lvdmaaten.github.io/tsne/

[5] UMAP: *Uniform Manifold Approximation and Projection for Dimension Reduction.* (s. f.). https://umap-learn.readthedocs.io/en/latest/

Photo. Kepler-11 System, NASA. https://www.nasa.gov/content/kepler-multimedia