509/E1

Machine learning application for binary and pulsation star classification

C.D. Anuruddha,* J. Adassuriya, and K. P. S. C. Jayaratne

[1]*Astronomy and Space Science Unit, Department of Physics*
University of Colombo, Colombo 03, Sri Lanka

In recent times, there have been many space missions exploring space. Some of these missions, like Kepler, K2, and TESS, focus on finding planets outside our solar system. But scientists are also interested in studying stars that change in brightness, called variable stars. These missions collected a lot of data, and researchers are using computer programs to analyse it. They use a special kind of computational technique called machine learning, which can quickly and accurately handle large amounts of data. The main goal of this research is to identify three groups of variable stars in the Kepler mission's data. These groups are pure pulsation stars, pure eclipsing binary stars, and eclipsing binary stars with pulsation components. To carry out our study, we downloaded 850 light curves from the MAST (Mikulski Archive for Space Telescopes) database and simulated 180 additional light curves using software that models binary stars. Here, light curves had to be simulated because there wasn't enough dataset for the class of pure binary stars. To classify the stars correctly, a machine learning algorithm called Random Forest was used. The reason for that was that it provided the highest accuracy among other algorithms used here (KNN, CNN, SVM and NB). The first classifier aimed to tell apart eclipsing binary stars and pulsation stars. The second one focused on categorizing four types of pulsation stars, *viz*., Delta Scuti, Gamma Doradus, RR Lyrae, and Solar-like Oscillation stars. Lastly, the third classifier was built to identify pure eclipsing binary stars and eclipsing binary stars with pulsation component systems. In this research 17 features were used and those were divided into two main categories, which are statistical features and Fourier analysis-based features. Here, 20% of the dataset was used to evaluate the first and second classifiers and 30% of the dataset was used to evaluate the third classifier. The first two classifiers showed an accuracy rate of 93%, which means they could distinguish between different types of pulsation stars and eclipsing binary stars very well. The third classifier achieved an accuracy rate of 89%, which is still quite good given the complexity of its task. This research is significant as it helps us classify the binary and pulsation stars in very large archives of Kepler and TESS. The classifier is available at the https://variable-star-classification-model.onrender.com/

Keywords: TESS data, machine learning, binary stars, pulsation stars

E-mail: anuruddhaedcd@gmail.com