MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

Contribution ID: **214** Type: **not specified**

Unsupervised Machine Learning Techniques for Young Stellar Object Light Curve Characterization

Tuesday 9 July 2024 15:50 (20 minutes)

Analysing light curve data presents unique challenges, notably the variability in data lengths, which complicates the use of traditional methods and many machine learning techniques designed for fixed-sized input vectors. In this study, we address these challenges by attempting to characterise Young Stellar Object (YSO) candidate light curves from the Gaia Photometric Science Alerts. The characterisation of YSO behaviour is essential to understand the underlying physics through the initial phases of star- and planet formation. While numerous statistics are available to describe temporal data, these are often less useful in the case of YSOs. Our approach leverages a Long Short-Term Memory (LSTM) based network, complemented by experiments with alternative architectures such as Transformers, to accommodate the variable length of our dataset. Central to our methodology is the implementation of an LSTM-based autoencoder designed to learn a compressed, fixed-size latent space representation of the light curves. This allows for the accurate reconstruction of time series from a uniform-dimensional embedded space, effectively converting variable-length data into fixed-length vector representations. To enhance our analysis, we combined the latent space generated by the autoencoder with statistical features for application in various clustering techniques, including k-means and density-based clustering methods such as DBSCAN. Subsequently, the downprojection of high-dimensional data to a 2-dimensional space using t-SNE was performed, enabling the identification of distinct YSO behaviour patterns. By using our method, we are able to identify a range of variability, including stochastic variability across different amplitudes, single or multiple outliers such as flares, dimming, and outbursting events on various timescales, as well as monotonic rising or fading.

Presenter: MADARÁSZ, Máté

Session Classification: Time domain