Contribution ID: **119** Type: **Plenary talk**

Using machine learning techniques in classification problems in Astrophysics

Tuesday, 6 September 2011 11:30 (40 minutes)

Multivariate datasets in astrophysics can be large, with the increasing volume of information now becoming available from a range of observations, from ground and Space, across the electromagnetic spectrum. The observations are in the form of raw images and/or spectra, and tables of derived quantities, obtained at multiple epochs in time. Large archives of images, spectra and catalogues are now being assembled into publicly-available databases: one example is the emerging global effort towards the Virtual Observatory. This necessitates the development of techniques that will allow fast, automated classification and extraction of key physical properties for very large datasets, and the ability to visualise the structure of highly multi-dimensional data, for extracting and studying substructures in a flexible way. Automated algorithms for clustering and outlier detection are necessary for a wide range of Astrophysical problems involving these growing datasets.
The applicability of commercial data mining tools is limited, since these do not incorporate the handling of errors in a principled manner, which is central to the analysis of Astronomical data, as it is in other branches of Physics. I will review how techniques used in the field of machine learning are being adapted for use in classification and clustering problems. Examples will include the use of topographic mapping to classify light curves of eclipsing binary stars, showing that this is an efficient way of searching for transiting extrasolar planets in large datasets, and robust density modelling for determining clusters and outliers, resulting in finding high-redshift quasars.

Primary author: Dr RAYCHAUDHURY, Somak (University of Birmingham)

Presenter: Dr RAYCHAUDHURY, Somak (University of Birmingham)

Session Classification: Tuesday 06th - Morning session

Track Classification: Track 2 : Data Analysis - Algorithms and Tools