HAMLET How to Apply Machine Learning to Experimental & Theoretical PHYSICS
August 19 - 21, 2024
Copenhagen, Denmark

Contribution ID: **26** Type: **Plenary**

Identifying dwarf AGN candidates through novel machine learning techniques

Tuesday 20 August 2024 13:55 (20 minutes)

While identification, characterisation, and triggering mechanisms of active galactic nuclei (AGN) have been since the 80's, the discussion has only been extended to include dwarf galaxies within the last decade.
This study aims to explore a novel AGN identification technique using a random forest (RF) classification technique, compare it to established identification methods, and investigate which set of properties/features constitute the best RF model.
Data is sourced from multiple catalogues: MaNGA (and its value added catalogue, Firefly) provides spatially resolved spectra of 10\,104 galaxies of which 1\,149 are dwarf galaxies. These galaxies constitute the base data set, and infrared (WISE) and X-ray (XMM) observations are matched to these. The NASA-Sloan Atlas is used for estimating environmental parameters.
The best model (from F1 score alone) is using internal features only of more massive galaxies. This model tends toward weighing fewer features higher and ignoring parameters that are less directly related to AGN ionisation. Conversely, this model disagrees the most with observations when it comes to dwarf galaxies, but provide twice as many dwarf AGN candidates as observations, and up to thrice as many compared to using intermediate mass galaxies as training set.
This approach provides a novel and interesting venue for identification of AGN in dwarf galaxies, but the method still requires fine tuning such as feature selection optimisation and validity assessment – are the predicted AGN actually AGN? If so, RF can be used to increase the sample size of known dwarf AGN and to adjust observational diagnostic diagrams in the low mass regime.

Primary author: KRISTENSEN, Mikkel Theiss (University of Southern Denmark)

Presenter: KRISTENSEN, Mikkel Theiss (University of Southern Denmark)

Session Classification: Parallel