Contribution ID : **65** Type : **Talk**

The AI Spectroscopist: Unveiling Galaxy and IGM Physics in the Era of Million-Spectrum Surveys

Monday, 7 September 2026 11:40 (30)

The era of wide-field spectroscopic surveys such as WEAVE, DESI, and 4MOST, along with forthcoming facilities like the Wide-Field Spectroscopic Telescope (WST) and MOSAIC, is delivering millions of spectra that encode the chemical and structural history of the Universe. The sheer volume of this data makes traditional human-supervised analysis intractable and necessitates the transition to fully automated machine learning (ML) pipelines to efficiently characterize the Intergalactic Medium (IGM), the circumgalactic medium (CGM), and complex galaxy and AGN physics.

A fundamental challenge in spectroscopic analysis is the accurate estimation of the intrinsic spectral continuum. We present an optimized autoencoder architecture that achieves superior precision, with a median Absolute Fractional Flux Error (AFFE) of 0.009 for quasar spectra. This model has demonstrated strong generalizability by successfully recovering the Lyα optical depth evolution in unseen DESI data.

Furthermore, specialized U-Net architectures are now capable of the generalized detection of metal absorption features (e.g., CIV, MgII, SiIV) with high completeness and purity, reaching an F1 score of ≈90% at S/N≈4. Following detection, physically motivated classifiers allow for robust ion identification with an average 90% accuracy, reaching ~100% for reliable doublet systems. This integrated pipeline allows the analysis of ≈10,000 spectra in just a few seconds.

Primary author(s) : PISTIS, Francesco (National Centre for Nuclear Research)

Presenter(s) : PISTIS, Francesco (National Centre for Nuclear Research)

Session Classification : Machine Learning in Astrophysics

Track Classification : Machine Learning in Physics