

Unsupervised classification of the 4XMM catalogue with variational autoencoders

More than 20 years of imaging the sky with XMM-Newton have produced hundreds of thousands of serendipitous sources. A systematic classification of these sources is far from trivial, given their diversity in nature, distance, brightness, and exposure, leaving us with a treasure trove that has yet to be fully opened. In this work, we perform an unsupervised classification of the 4XMM catalogue using a variational autoencoder (VAE) approach applied to the EPIC spectral information of each detection. This method helps us cope with the large dynamic range and noise levels that characterise the catalogue. We show that working with the sources in the latent space facilitates the discovery of objects of interest and reveals their dynamical evolution through latent trajectories. We further explore the use of semi-supervised clustering to enhance the efficiency of the autoencoder. Such approaches can be readily extended to the upcoming 5XMM catalogue and contribute to the identification of targets of opportunity.