Abstract: Spectroscopy has been utilised in identifying the main power source in active galaxies. Based on the different mechanisms that excite the gas that exists inside the galaxies (and which, as a result of those mechanisms, glows in different wavelengths), the Galaxies may be separated into 4 categories: the AGN (Active Galactic Nuclei), which are divided into LINERs and Seyferts, the HII region-like galaxies(star forming galaxies) and the Composite galaxies (their spectra contain significant contributions from both AGN and star-forming).

Four emissions intensities ratios are being used as means to classify different galaxies; log(NII/Hα), log(SII/Hα), log(OI/Hα) and log(OIII/Hβ). Both physical and empirical models have been developed in order to propose a classifcation scheme based on those ratios. However, the exact demarcation between star-forming galaxies and AGN is subject to considerable uncertainty and the increasing flow of data from massive new surveys shows the inadequacy of the existing scheme.

In this project we utilise a data-driven approach in order to build a density estimation model that will describe accurately the distributions of the 4 different classes of galaxies. Identifying and parametrizing the distributions of the pure star-forming Galaxies and the pure AGN would provide a solid quantitative tool in order to explore further scientific problems.