

A Machine Learning Method to Achieve High Accuracy in Galaxy and AGN Classification using Photometric Data

Sushanth Elangovan, Antonio C. Rodriguez

SUPPORTING INFORMATION

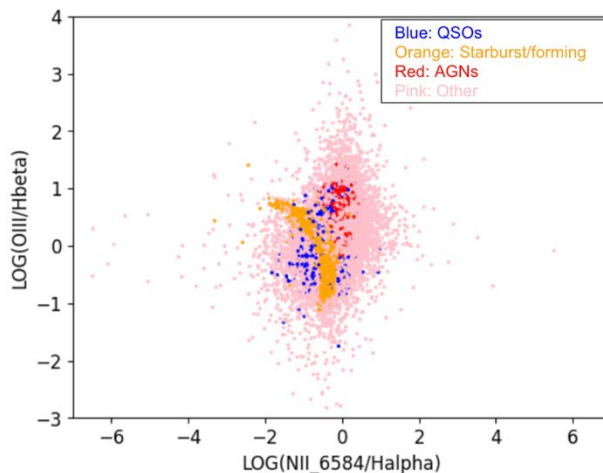


Figure S1. Original BPT diagram plotted into four groups: blue (QSOs), orange (Starburst/forming), red (AGNs), pink (other galaxy types). The distribution and clustering of QSOs, Starburst/forming, and AGNs can be seen on spectroscopic axes $\text{LOG}(\text{OIII}/\text{Hbeta})$ and $\text{LOG}(\text{NII}_{6584}/\text{Halpha})$. Misclassification is noticeable as datapoints overlap across multiple categories, leading to the need for a better classification model to clearly separate the groups.

Table S1. Random Forest Classification of AGNs within filtered GALAXY-only dataset (photometry and spectroscopy readings)

n_estimators	test_size	accuracy
64	0.15	0.81
64	0.2	0.81
64	0.25	0.81
100	0.1	0.81
100	0.15	0.81
100	0.2	0.81
100	0.25	0.82

Table S2. Logistic Regression of AGNs within filtered GALAXY-only dataset (photometry and spectroscopy readings)

test_size	accuracy
0.1	0.76
0.15	0.77
0.2	0.77
0.25	0.77