



TIMESTEP Summer Internship with NOIRLab

Rohin Sant



What is NOIRLab?

NSF funded, US centre for ground-based, night-time optical astronomy. Some important programs under NOIRLab:

- Vera C. Rubin Observatory.
- Kitt Peak National Observatory.
- International Gemini Observatory.
- Many important observatories and science programs such as Astro Data Lab.



Iyer et al. 2024 ("Pathfi")

Joel Shernicoff & Rohin Sant

Astro Data Lab

Astro Data Lab enables astronomers and the public to use and analyse large datasets produced by instruments at NOIRLab.

- It has a collection of 75 Jupyter Notebooks.
- Tutorials and Science Cases.
- Contains data from SDSS and DESI.
- A Notebook Gallery would enhance the delivery of the work done at NOIRLab.

Astro Data Lab Science Platform

The Astro Data Lab Science Platform enables efficient exploration and analysis of the large datasets now being generated by instruments on NOIRLab and various other wide-field telescopes.

New to Data Lab?

To learn more about the Astro Data Lab Science Platform click the "Getting Started" link below. This page provides detailed information about the platform including how to register an account and how to use the notebook server, client libraries and data explorer.

[Getting Started >](#)

[Existing Users Sign In Here >](#)

[Don't have an account?](#)
You can also [explore Data Lab without signing in](#).

Image credit: Astro Data Lab

Iyer et al. 2024 ("Pathfi

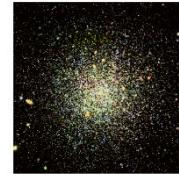
Joel Shernicoff & Rohin Sant



Goals of my projects

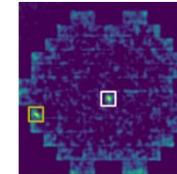
- To create a notebook gallery that helps users to discover, filter and launch Astro Data Lab Jupyter Notebooks.
- To use machine learning techniques to study quasar/AGN activity in nearby galaxies.

Astro Data Lab Notebook Gallery



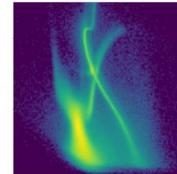
[Getting Started with Data Lab](#)

Learn the basics such as importing modules, sending a database query, and using the Simple Image Access (SIA) service to create image cutouts.



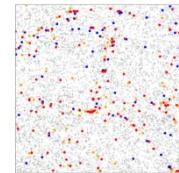
[Dwarf Galaxies in the SMASH survey](#)

Discover the ultrafaint Hydra II dwarf galaxy in the SMASH DECam survey based on spatial overdensities of blue stars with a detection algorithm.



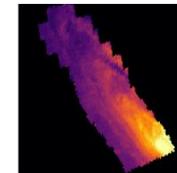
[Exploring SMASH DR2](#)

The 480 square degree SMASH DECam survey of the Magellanic Clouds and their periphery contains a wealth of objects, including this capture of the SMC with 47 Tuc in the foreground.



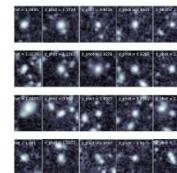
[Large-Scale Structure of the Universe](#)

Investigate cosmic filaments and clusters of galaxies, pan around an interactive sky viewer, combining spectroscopy and DESI pre-imaging.



[Fun with PHAT](#)

Visualize the 100 million+ stars in the Andromeda Galaxy captured by the Panchromatic Hubble Andromeda Treasury (PHAT).



[Gallery of Cluster Galaxies](#)

Use the Simple Image Access (SIA) service to retrieve images from the Gemini GOGREEN program.

Image credit: Data Lab Team

Joel Shernicoff & Rohin Sant



Notebook Gallery Project

- Skim-read Jupyter Notebooks, the README file and keyword file in the *notebooks-latest* GitHub repo.
- Made a pros and cons list with our recommendations for the gallery.
- We explored external galleries for comparison.
- Compared Data Lab's Keywords with the Unified Astronomy Thesaurus (UAT).
- Presented results and gathered input from Data Lab members.
- Made a visual of suggested gallery with our recommendations.



notebooks-latest

Public

01_GettingStartedWithDataLab

02_DataAccessOverview

03_ScienceExamples

04_HowTos

05_Contrib

06_EPO

tests

.gitignore

CONTRIBUTING

DataLabNotebookTemplate.html

DataLabNotebookTemplate.ipynb

LICENSE

README.txt

keywords.txt

Image credit: Data Lab Team



Joel Shernicoff & Rohin Sant



Notebook Gallery Recommendations

- Add difficulty levels for audiences.
- Add search options, next and previous notebooks.
- We explored external galleries for comparison.
- UAT Keywords as search options with difficulty tags.
- Sidebar navigation and dark mode.
- Sort keywords based on target search.
Example: science case, etc.

Jupyter Notebook Gallery

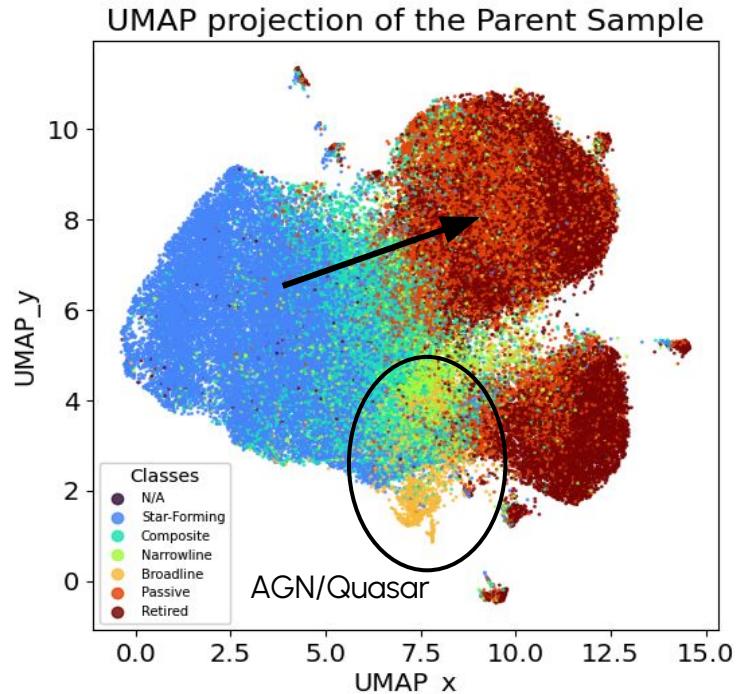
Tutorials	Science Cases
> Python Basicss	> Galaxies
> SQL Queries	> White Dwarfs
> NumPy	> Star Clusters
> Matplotlib	

Variables & Data Types Introduction to variables, Integers, strings in Python	Start
Loops and Functions For loops, while loops and defining functions	Start
Using NumPy with FITS Files Open FITS images using NumPy arrays	Start
Galaxy Classification	Start



Research Project

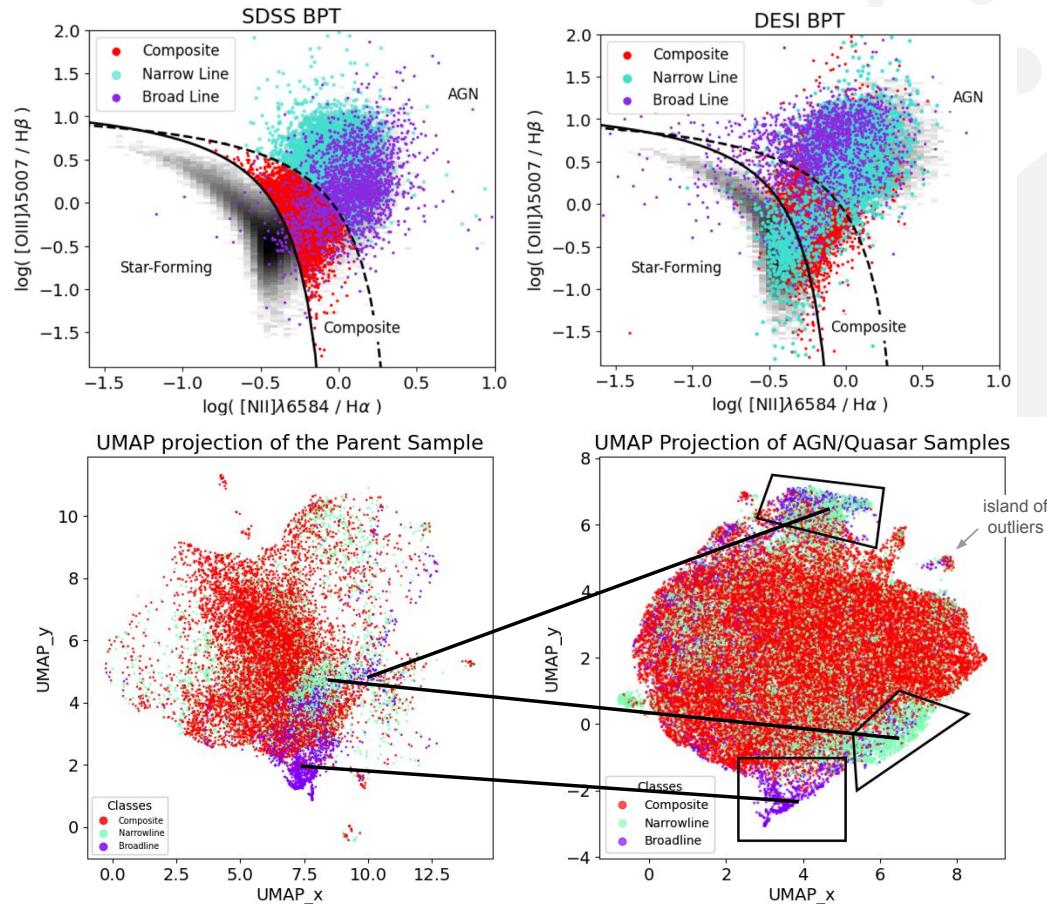
- Worked with Dr. Stéphanie Juneau to use machine learning to see AGN/Quasar activities.
- Used datasets from Astro Data Lab.
- AGN's are powered by black holes and they influence the host galaxies.
- Spectra of galaxies encode information of multiple physical properties.
- UMAP is a ML algorithm to visualize this data.
- Pat et al. (2020) used UMAP on low-redshift galaxies to visualize trends.





Some Findings

- Using emission-line diagnostics, created BPT diagrams.
- Some discrepancies seen in BPT's.
- Applied UMAP on just the AGN/Quasar samples of the parent dataset.
- We see regions correspond to each other on UMAP's.

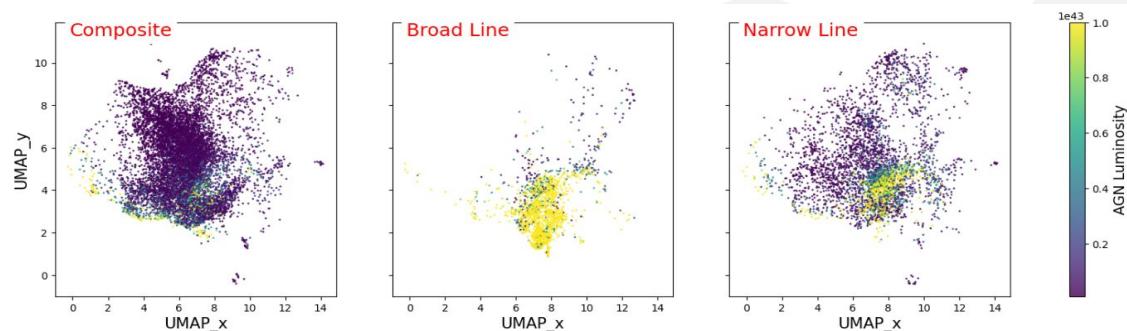




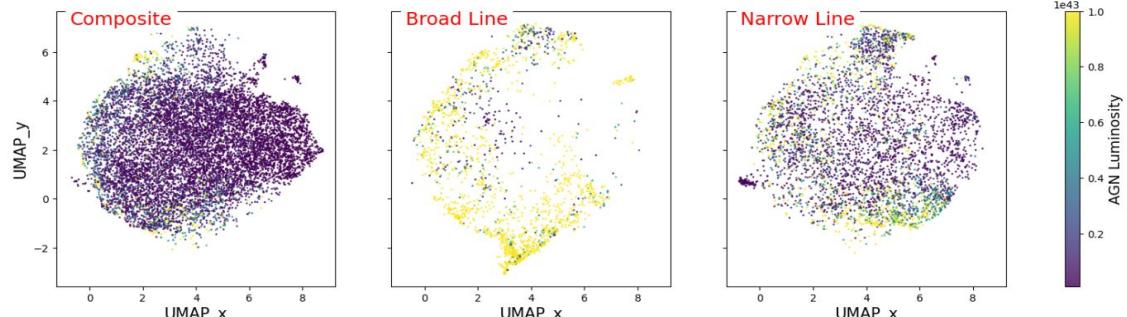
More Findings

- Color-coded plots with line ratios and AGN Luminosity.
- Most AGN luminous galaxies present in broad-line AGN as quasars accrete a lot of material around supermassive black holes.
- The most luminous narrow-lines and composites are near broad-lines.

Top row: Parent sample
UMAP split into AGN
classes



Bottom row: AGN-only
UMAP split into AGN
classes





Summary and Skills Learnt

- Created a complete review of Astro Data Lab current notebook collection.
- Conducted state-of-the-art research about AGN/Quasar trends and activities using machine learning and Astro Data Lab.
- Presented my research work in a poster with Dr. Juneau at an astronomical software conference.
- Met with various scientists and had an amazing time overall!
- Skills learnt:
 - Jupyter Notebooks
 - SQL Queries
 - Presentation Skills
 - Using ML techniques



Thank you!

Acknowledgments

This work was supported by the NSF-Simons CosmicAI institute, the University of Arizona's TIMESTEP program and made use of data and services from the NSF NOIRLab Astro Data Lab. This work is based in part on data from the Sloan Digital Sky Survey (SDSS) and the Department of Energy Dark Energy Spectroscopic Instrument (DESI).