An updated syllabus can be found here.

Office Hours:

TF office hours are Mondays 4:30-5:30 in A-101 (at the Center for Astrophysics)

Dr. V's office hours are Wednesdays 12:00-1:00 in P-243. Dr. V will also hold office hours at 4pm on Fridays by appointment only.

Course goals:

In this course, students will (1) become familiar with a wide range of data-driven techniques, (2) be able to utilize domain knowledge in the development of data-driven pipelines and (3) implement data-driven techniques for astrophysical applications in Python.

Note that this course is being offered at the undergraduate level (as a Senior-level course) and the graduate level. Undergraduates will be required to have background knowledge in linear algebra (MATH 21b or equivalent) and astrophysics (ASTRO 16 and ASTRO 17 or equivalents). Graduate students have no prerequisite courses. Both undergraduate and graduate students are strongly encouraged to have experience with Python. Students are not expected to have any background in data science.

Students outside of astronomy are welcome to enroll. However, course problem sets will primarily pull from the astrophysical literature.

Course Summary

Date	**Topic**	**Helpful Readings**
Tuesday, Jan 23	Introduction to MOO	Murphy 1.1-1.3
Thursday, Jan 25	Introduction to MOO, Probability	Murphy 2.1, 2.2, 2.4.1
Tuesday, Jan 30	Introduction to MOO, Probability	
Thursday, Feb 1	Multi-layer perceptrons	Murphy 16.5
Tuesday, Feb 6	Multi-layer perceptrons	
Thursday, Feb 8	PCA, t-SNE	Murphy 12.2, 12.3; distill.pub
Tuesday, Feb 13	t-SNE, MLP-based autoencoders	Murphy 28.3
Thursday, Feb 15	k-means (EM algorithm), Gaussian mixture models	Murphy 11.1, 11.4
Tuesday, Feb 20	Gaussian mixture models, self-organizing maps	Murphy 11.1, 11.4
Thursday, Feb 22	Self-organizing maps, DBScan	Kind & Brunner 2014
Tuesday, Feb 27	Decision trees	Murphy 16.2
Thursday, Feb 29	Random forests	Murphy 16.2
Tuesday, Mar 5	Random forests	Murphy 16.4(.3 for AdaBoost)
Thursday, Mar 7	Probabilistic random forests	Reis, Baron & Shahaf 2019
Tuesday, Mar 19	MLP (review), convolutional NNs	Murphy "Book 1", 14.1, 14.2
Thursday, Mar 21	CNN, Recurrent NNs	Murphy Book 1, 15.2
Tuesday, Mar 26	RNNs, Graph NNs	Villanueva-Domingo et al
Thursday, Mar 28	GNNs	
Tuesday, Apr 2	Dropout; coding examples	Murphy Book 1, 13.5.4
Thursday, Apr 4	Review of Bayesian Statistics	Murphy 2.1, 2.2, 2.4.1
Tuesday, Apr 9	Variational Methods	Blei, Kucukelbir & McAuliffe (2016)
Thursday, Apr 11	Variational AEs	
Tuesday, Apr 16	Diffusion Models	Luo (2022)

Thursday, Apr 18	Normalizing flows and Simulation-based inference	Rezende & Mohamed (2015)
Tuesday, Apr 23	Poster Session	

Note: "Murphy" corresponds to "Book 0" in the Machine Learning series. The series additionally includes Book 1 and Book 2 (drafts available online).