ASTR 324: Introduction to Astrostatistics and Machine Learning in Astronomy

Mario Jurić and Stephen Portillo

University of Washington, Spring Quarter 2021

Location and Time: TTh 10:00am-11:20am, **Zoom at** `http://dirac.us/324`

Office Hours: After Thursday class, via Zoom

Grading: homeworks, 70%; final exam: 20%; quizes: 10%.

Class materials: `https://github.com/uw-astr-324/astr-324-s21`

Class JupyterHub: `https://dirac.us/hub324`

UW Astronomy Slack: `https://join.slack.com/t/uw-astronomy/signup`, then join `#astr324`

Textbook: Ivezić, Connolly, VanderPlas & Gray: *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*

Flipped classroom with online teaching:

This course will follow the flipped classroom model. In this method of teaching, you will listen to (prerecorded) lectures at home, and come to class (virtually, via Zoom) to engage in discussion, group work, and work on homeworks.

A typical week will run as follows:

- **by Friday 6pm:** lectures for next week published on class YouTube channel. Please review them (as well as supplement them with chapters from the textbook) in time for a short quiz and survey on Monday. New homework will be assigned at the same time.
- **by Monday 5pm:** take the quiz, fill out the survey about what was not clear.
- **Tuesday and Thursday class**: group discussion about outstanding questions, work on homeworks. Ideally, you will have finished your homework by the end of Thursday class.
- **by Friday 6pm:** lectures (and homeworks) for next week posted, and the cycle repeats...
- **by 6pm the *following* Friday:** homework assigned two weeks ago are due.

Learning Goals:

This course introduces students to statistical and computer science tools and techniques commonly used in data driven astronomy and astrophysics. It does so through a hands-on approach, with theory followed by working through examples of data analysis with modern astronomical datasets. Practical data analysis is done using Python tools, with emphasis on the astroML module (see www.astroML.org). The lectures taught at the undergraduate level, designed for astronomy and physics majors. The main discussion topics are based on Chapters 4 and 5, and selected topics from Chapters 6-10, from the reference textbook.

Prerequisites:

Students taking this class are required to have basic calculus and basic Python skills, as well as basic scientific measurements and statistics skills at the level of a freshman lab.

Topics:

- WEEK 1 (starting Mar 30): Getting started with online/flipped classroom learning
- WEEK 2 (starting Apr 6): Introduction to probability and statistics I
- WEEK 3 (starting Apr 13): Introduction to statistics II
- WEEK 4 (starting Apr 20): Maximum likelihood and applications in astronomy
- WEEK 5 (starting Apr 27): Bayesian inference and model selection
- WEEK 6 (starting May 04): Markov chain Monte Carlo methods
- WEEK 7 (starting May 11): Dimensionality reduction
- WEEK 8 (starting May 18): Time series analysis
- WEEK 9 (starting May 25): Introduction to Machine Learning I
- WEEK 10 (starting June 01): Introduction to Machine Learning II
- FINAL EXAM: June 7 (to be submitted by 5pm): take-home, limited-time, closed book final exam.

Homeworks:

There will typically be a homework each week, assigned on Friday (6pm), and due **two weeks later** on Friday (6pm). The homeworks will be centered on practical work using Python, designed to practice what we've learned in the week after the homework is assigned. All homeworks will involved writing Jupyter notebooks.

Quizzes:

Multiple-choice quizzes will be every due every Monday afternoon (5pm), about key concepts covered in lectures posted the previous Friday.

Final Exam:

The final exam will consist of simple questions with few-sentence answers, asking about the key concepts we've learned about and discussed in class. It will be a "take home", closed book, and limited time exam, on an honor system.

Timeliness policy:

- All homeworks are due **two weeks** after being assigned.
- Homeworks turned in within one week after posting will gain a 10% point bonus.
- Homeworks turned in late (up to a week) will receive a 20% point deduction.
- **Homeworks turned in more than a week late receive a 50% point deduction.**.

Privacy/FERPA statement:

This course is scheduled to run synchronously at your scheduled class time via Zoom. These Zoom class sessions **may** be recorded. The recording will capture the presenter's audio, video and computer screen. Student audio and video will be recorded if they share their computer audio and video during the recorded session. The recordings will only be accessible to students enrolled in the course to review materials. These recordings will not be shared with or accessible to the public.

The University and Zoom have FERPA-compliant agreements in place to protect the security and privacy of UW Zoom accounts. Students who do not wish to be recorded should:

Change their Zoom screen name to hide any personal identifying information such as their name or UW Net ID, and Not share their computer audio or video during their Zoom sessions.