CBIO 4835/6835

Introduction to Computational Biology

Biological Sciences, Room 404A
Tuesdays & Thursdays, 9:30-10:45am

Dr. Shannon Quinn

Email: squinn@cs.uga.edu
Website: https://eds-uga.github.io/cbio4835-fa18
Office Location: Boyd GSRC 638A
Office Hours: TBA

The course syllabus is a general plan; when (not if) deviations arise, they will be announced.

Course Description: Computational Biology, a highly relevant and fast growing subfield in biology, is an interdisciplinary effort to bring computer science, machine learning, and data mining techniques to the wet lab environment, automating experiments and providing objective, quantitative results. We will examine fundamental use cases in biology such as bioimaging, genomics, structure analysis, dynamic systems modeling, and other topics in systems biology through the lens of the Python programming language and its ecosystem. A diversity of analytical techniques will be described in detail.

Prerequisites: None required; you should have an understanding of basic concepts in biology and biochemistry, as we will explore these through the lens of programming in the course. It is fully expected that few, if any, will have substantive prior programming experience. **If you have substantial prior programming experience, this class is not for you.**

Credit Hours: 3

Text(s): We will have **no** required text book for this class. Recommended textbooks include:

1. *Bioinformatics Programming Using Python* (1st ed.)
   https://www.amazon.com/Bioinformatics-Programming-Using-Python-Biological/dp/059615450X/.
2. *Bioinformatics Data Skills* (1st ed.)
   http://shop.oreilly.com/product/0636920030157.do.
3. *Building Machine Learning Systems with Python* (1st ed.)
   https://www.amazon.com/Building-Machine-Learning-Systems-Python/dp/1782161406/.

Topical Course Outline

This course will be split into roughly four modules, interleaving the fundamentals of programming in Python with their application in the four focus areas of computational biology. Relevant Python packages will be discussed and explored, and the assignments will provide an opportunity for hands-on interaction with data.

Computational Genomics

- Sequence alignment
- Dynamic programming
- Hidden Markov Models
- Command line; Python syntax and variables; loops; data structures; conditionals

Computational Modeling

- Compartment models
- Differential equations
- Agent-based modeling
- NumPy arrays, functions, probability, statistics, biopython, bioconda

Structural biology

- Coordinate systems
- Molecular dynamics simulations
- Clustering and segmentation
- PyMol, scikit-learn, subprocess

Bioimaging

- Computer vision
- Object segmentation and tracking
- Dynamic textures
- scikit-image, OpenCV, matplotlib

Grade Distribution:

Students in 6835

Participation	5%
Assignments	45%
Midterm Exam	15%
Final Project	15%
Final Exam	20%

Students in 4835

Participation	5%
Assignments	60%
Midterm Exam	15%
Final Exam	20%

There will be 6 assignments, each two weeks long, and worth either 10% each (4835) or 9% (6835). Students who do the final project are exempt from doing the 6th and final assignment; this project is required for students in 6835 but optional and worth extra credit for students in 4835. The final project will consist of several parts, graded individually and aggregated to 15% of your final grade. Midterm and final exam layouts will be discussed closer to the exam dates. Participation can be earned in several ways, including but not limited to 1) answering questions and participating in

discussions on Slack, 2) attending office hours, and/or 3) engaging in lecture.

Course Policies

- **Attendance**
  - Come to lecture. If you're not in lecture to answer every question I pose to the class, and you have a bad day on the midterm, I won't have any way to know that you actually understand the material. This course has all of 20 people registered; I'll know when someone is missing.
  - If you cannot attend lecture, let me know ahead of time and we'll work something out.
  - Try to keep in-lecture hacking on your laptop to a minimum. I certainly understand testing out methods that we discuss in class, but I may interleave critical exam hints into the lecture if I detect a lack of attentiveness; I cannot be held responsible if these hints fall on distracted ears.
- **Assignments**
  - There will be six 2-week assignments, focusing on using Python to answer certain biological questions.
  - Assignments are released and submitted through JupyterHub. Contact me immediately if you are unable to access JupyterHub for any reason, but keep in mind that you can only access the site either from the UGA campus network or while connected to the VPN.
  - Assignments are due by 11:59:59pm on the noted date. Assignments turned in after this deadline will lose 25/100 points for every subsequent 24 hour-period they are late.
  - All assignments are to be completed individually, but you are welcome and encouraged to collaborate with other students when conceptualizing and framing the problems. You can also set up Slack channels to do this!
  - *The presence or absence of any form of help or collaboration, whether given or received, must be explicitly stated and disclosed in full by all involved, on the first page of their assignment* ("I did not give or receive any help on this assignment" or "I helped [person] with [specific task]."). Collaboration without full disclosure will be handled severely; except in usual extenuating circumstances, my policy is to fail the student(s) for the entire course.
  - It's really easy to tell if code has been copied. Coding styles are like handwriting; everybody's is unique.
- **Final Project**

  Graduate students are **required** to do a final project; for undergraduates, it is **optional** and worth extra credit. The final project consists of three main deliverables: a proposal outlining your project goals, a 25-minute presentation on your results the last day of lecture, and a written conference-style paper detailing your problem statement, methods, and findings. If you do the final project, you do NOT have to do the 6th and final assignment.

  To repeat: those who do a final project are exempt from Assignment 6. The final project is required of graduate students, and optional for undergraduates (worth a full letter grade of extra credit).

Academic Honesty

As a University of Georgia student, you have agreed to abide by the University's academic honesty policy, "A Culture of Honesty," and the Student Honor Code. All academic work must meet the standards described in "A Culture of Honesty" found at: `https://ovpi.uga.edu/academic-honesty/academic-honesty-policy`. Lack of knowledge of the academic honesty policy is not a reasonable explanation for a violation. Questions related to course assignments and the academic honesty policy should be directed to the instructor.

- Read "A Culture of Honesty," the UGA academic honesty policies and CS Academic Integrity policies.
- You must not allow others to copy or look at your work.
- You must not give/share your lab/project assignment work to a fellow student.
- Copying significant portions of code from a fellow student or any other source (including internet) is plagiarism and will be dealt with as such.
- If you have questions about an assignment or if you run into problems, contact your instructor/lab instructor.
- During exams, no assistance and no additional materials are allowed.
- All of your coursework must meet the aforementioned policies and rules. Students that violate any of these rules or the UGA Academic Honesty policies will be liable to a penalty. The instructor will strictly enforce Academic Honesty policies and report any violation of the aforementioned policies and rules.