PROGRAMMING FOR BIOINFORMATICS – BIOL 7200 – Syllabus – Fall 2021

Instructor: **Lavanya Rishishwar (**lavanya.rishishwar@gatech.edu**)**
Teaching Assistants: **Ujani Hazra (**ujani.hazra@gatech.edu**),**
Shivam Sharma (ssharma454@gatech.edu**)**

Course summary: The fields of Bioinformatics and Computational Biology occupy the intersection of the life sciences and information technology. Over the last decade, there has been an explosion of data in the life sciences, and the proliferation of raw information promises to continue at an even more rapid pace. Computers are needed to handle and assimilate this massive amount of information. More importantly, the role of bioinformatics is to convert information, in the form of data, into biological knowledge. In order to do this, bioinformaticists and/or computational biologists must be adept at the use of computers, *i.e.* YOU MUST KNOW HOW TO CODE.

This project-based/lab course will introduce programming for bioinformatics. We will begin by introducing you to the command line environment in the Unix / Linux operating system – this is where real scientific computing gets done. This will include a broad coverage of Unix / Linux utilities as well as shell scripting. The course will then go on to use the Python programming language to illustrate the fundamentals of bioinformatics programming.

All required and recommended readings, lectures and exercises will be made available on the course Canvas site (https://canvas.gatech.edu/). This is an exclusively practical and active learning class. Students will complete exercises in order to learn how to code and how to do bioinformatics. The only way to learn the course material is by doing. Students will have the opportunity to demonstrate and explain their code to the class. Students will be required to post exercise answers and codes to the course Canvas site for evaluation. Assignments are due by 11:59 PM on Mondays and must be posted to Canvas by that time. Late assignments will not be accepted and will result in a score of 0. Thursday follow-up programmatic discussion sessions will serve as code demonstration and lab sessions where students will get the opportunity to discuss class problems with the TAs. Office hours can be requested by email.

Please see http://osi.gatech.edu/content/honor-code for Georgia Tech's Academic Honor Code, which you are required to uphold.

Diversity and Disability Statement

Georgia Tech values diversity and inclusion; we are committed to a climate of mutual respect and full participation. Our goal is to create learning environments that are usable, equitable, inclusive and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please notify the instructor as soon as possible. Students with disabilities should contact the Office of Disability Services to discuss options of removing barriers in this course, including accommodations. ODS can be reached at 404.894.2563, dsinfo@gatech.edu, or disabilityservices.gatech.edu

Course Evaluation:

Demos	25%
Code/exercise submission & evaluation	75%
• Code/assignment submission	*10%*
• Quizzes	*10%*
• Code evaluation	*55%*

Schedule of lecture / lab sessions

Date	Topic	Room
8/24/2021	Introduction to *nix environment	Mason 2117
8/26/2021	Setup of *nix environment	Mason 2117
8/31/2021	Basic of system administration and software installations	Mason 2117
9/2/2021	Downloading datasets from internet	Mason 2117
9/7/2021	Regex (sed, grep)	Mason 2117
9/9/2021	Follow-up programmatic discussion	Mason 2117
9/14/2021	Programming with AWK	Mason 2117
9/16/2021	Follow-up programmatic discussion	Mason 2117
9/21/2021	Introduction to version control and Bash scripting	Mason 2117
9/23/2021	Follow-up programmatic discussion	Mason 2117
9/28/2021	Creating Bash based pipelines	Mason 2117
9/30/2021	Follow-up programmatic discussion	Mason 2117
10/5/2021	Introduction to Python	Mason 2117
10/7/2021	Follow-up programmatic discussion	Mason 2117
10/12/2021	*Fall break*	
10/14/2021	Multidimensional data structures	Mason 2117
10/19/2021	Commandline arguments and logging	Mason 2117
10/21/2021	Follow-up programmatic discussion	Mason 2117
10/26/2021	Recursion and dynamic programming: sequence alignment	Mason 2117
10/28/2021	Follow-up programmatic discussion	Mason 2117
11/2/2021	Working with common bioinformatics formats	Mason 2117
11/4/2021	Follow-up programmatic discussion	Mason 2117
11/9/2021	Working with genomic intervals	Mason 2117
11/11/2021	Follow-up programmatic discussion	Mason 2117
11/16/2021	Modules and parallel computing	Mason 2117
11/18/2021	Follow-up programmatic discussion	Mason 2117
11/23/2021	Review of concepts covered in the class	Mason 2117
11/25/2021	*Thanksgiving break*	
11/30/2021	Review of concepts covered in the class	Mason 2117
12/2/2021	Follow-up programmatic discussion	Mason 2117
12/7/2021	Final instructional Day	Mason 2117

Note that the syllabus is subject to change depending on how the pandemic scenario goes, the speed at which the class progresses and the performance of the students.