BIOI 500 Advanced Bioinformatics

- Instructor: Dr. Heather Wheeler hwheeler1@luc.edu (use hewlab15@gmail.com if sending a .py file)
- Prerequisites: BIOL 388/488
- Days/Times: Tues/Thurs 1:00 – 2:15 PM
- Classroom: IC 111 (starting Feb 1) or Zoom (Jan 18-27, see Zoom Pro tab on Sakai for link)
- Instructor Office Hours: Class work time or by appointment

Objectives:

- Describe bioinformatics algorithms and tools for sequence analysis
- Reproduce and understand solutions for sequence analysis
- Analyze literature in the field of bioinformatics
- Design, implement, and evaluate algorithms in a topic area relevant to bioinformatics

Course Materials: The textbook *Bioinformatics Algorithms* by Drs. Compeau and Pevzner is the starting material for many topics covered in class and is strongly recommended. An online version is available at https://www.bioinformaticsalgorithms.org/read-the-book (first 5 chapters free) or you can purchase/rent a paper copy (ISBN: 0990374637). We will also use free online materials and the scientific literature. Students must have access to a computer (Mac, Windows, or Linux operating system) with administrative rights, i.e., can install software. Computers must have a minimum of 8 GB of RAM and a reliable internet connection. Students will be given an account to access a remote Linux server if needed for completing the final project. Homework assignments require Python and can be completed on a personal computer. For synchronous Zoom sessions, the student needs a device with a microphone and a large enough screen to view shared code. A webcam is helpful for interacting with classmates and instructor.

Seeking Help: To meet at a time other than scheduled class time, you must email the instructor to schedule a day and time in advance. Note that many workdays are built into the course schedule where you will have ample time to work on assignments and ask questions. I recommend you attempt homework assignments before workdays so you can get your questions answered during class time. Another route to get assistance is to ask questions via email. I will do my best to respond to emails within 24 hours. There are many online forums where you can post your questions or search for questions and answers. When you encounter a problem, error message, etc., you can always try and **Google it!**

Student Accommodations: Any student requesting accommodations is required to register with Student Accessibility Center (SAC). Students will provide the instructor with an accommodation notification from SAC, preferably within the first two weeks of class. Students are encouraged to meet with the instructor individually in order to discuss their accommodations. All information will remain confidential. For more information or further assistance, please call 773.508.3700 or visit https://www.luc.edu/sac/.

Academic Integrity: Plagiarism and cheating will not be tolerated and will be reported to the college. See http://www.luc.edu/academics/catalog/undergrad/reg_academicintegrity.shtml. STUDENTS FOUND TO PRESENT SOMEONE ELSE'S WORK AS THEIR OWN WILL **RECEIVE ZERO POINTS FOR THE ASSIGNMENT**. ANY STUDENT WHO REPEATS SUCH AN ACTION WILL RECEIVE A FAILING GRADE (F) FOR THE COURSE.

The programming and project assignments in this course will be done independently (although you can get help from each other in class). Cheating includes submitting as your own work

something that has been written by another person (classmate, tutor, website, etc.). **You must comment your code in your own words** to demonstrate that you understand it. You will often be asked questions about how your code works and your comments will help you explain. **Inability to explain how your code works will result in a reduced grade.**

Your writing assignments must properly cite the literature. The scientific literature typically does not use direct quotations, so you should not use them in this course. When citing the work of others work, simply changing a word or two in a sentence is plagiarism. If you are concerned that your text may boarder on plagiarism, ask. ***If you are found to present someone else's work as your own, you will receive zero points for the assignment. If you present someone else's work as your own a second time, you will receive an F for the course.***

Grading: Your grade will comprise several assignments as listed below.

Assignment	Points
ROSALIND Homework	40
Journal Club Presentation	15
Journal Club Participation	18
Final Presentation	10
Final Presentation Participation	2
Final Paper	25

- **ROSALIND Homework: 40 pts**. We will use the ROSALIND web platform to solve assigned Python programming problems that implement bioinformatics algorithms. On the due date in class, you may be chosen for a code check. If chosen, you will be asked questions about how your code works. Inability to answer such questions will result in a reduced grade. Our goal is for you to apply Python programming to better understand algorithms used in bioinformatics, not for you to simply solve the ROSALIND problems without understanding your code.
- **Journal Club Presentation: 15 pts**. Your journal club presentation will relate to your final project. For this presentation, you will select one paper that describes one of the three approaches/tools you will test in your final project. Choose a paper published within the past 5 years to present. Sign up for the day you want to present at this link, first come, first served: https://doodle.com/poll/skcq4mvpzgub4h6u?utm_source=poll&utm_medium=link. You must submit the paper you choose to present via email to Dr. Wheeler at least one week prior to your scheduled presentation so that it can be made available to your classmates via Sakai. Your presentation will briefly describe the problem that the approach/tool is designed to solve. This will require presenting background information about the problem that may not be included in the paper itself, i.e., how would you introduce this problem if you were teaching a class about it? Although you will choose one paper/tool to focus on at journal club, you may include background information about the other two tools you are testing in your presentation. As the presenter, you will then engage your classmates in a discussion of the paper. This will likely include asking classmates to describe a figure or table, posing thought questions about the approach and its limitations, answering questions posed by the group or directing them to others, etc. This presentation and discussion is expected to last 45-60 min. A rubric of expectations will be posted on Sakai. The journal club presentation on February 10 will serve as an example presentation. Your presentation slides (pptx/pdf) must be submitted via Sakai Assignments prior to the scheduled presentation.

- **Journal Club Participation: 18 pts**. When someone other than yourself is the journal club presenter, you are expected to read the presented paper before class and actively participate in the discussion of the paper during class. If you make at least one substantive comment (e.g., describe a figure or table, explain a method or result, answer a thought question posed by the presenter, ask an insightful question, etc.), you will be awarded 2 pts. It is possible to miss participating in 1 journal club and still receive the maximum 18 points.
- **Final Project:** You will work independently to evaluate and benchmark 3 computational tools. Benchmarking projects started in this class have led to publications, for example https://doi.org/10.7717/peerj.10090. By March 3, you must discuss your proposed project with Dr. Wheeler (this can occur during class workdays).
  - **Final Project Presentation**: **10 pts**. This presentation will be an 8-minute talk with 3 minutes for Q&A. The presentations should introduce the problem, briefly introduce the tools evaluated (highlighting differences/similarities in their approaches), the results of your tests, and a conclusion (i.e. when should you use tool 1?, when does tool 2 fail?, which tool is easiest to use?, etc.) A grading rubric will be posted on Sakai. The presentation document (pptx/pdf) must be submitted via Sakai prior to the scheduled presentation.
  - **Final Project Presentation Participation**: **2 pts**. You can earn up to 2 points for asking questions of other student presentations (1 pt/question). Questions should focus on interpretation of the results or aspects of the tool (rather than simply asking a presenter to repeat/reshow a slide.) Note, time is limited for Q&A. Everyone will not have an opportunity to ask questions of each presentation.
  - **Final Project Paper: 25 pts.** Final project topics will be determined by March 4. The final project will not include new code development. Rather, you will evaluate publicly available solutions and review/benchmark the tools. This analysis will be reported through a manuscript including the following sections (a grading rubric will be posted on Sakai):

    Abstract. 200-300 word summary of the paper.
    Introduction. Introduce the biological problem examined. Briefly describe available tools (including landmark historical approaches, or other variations). Introduce the approach taken by each of the three tools you investigated.
    Methods. Include details about the data used to benchmark the tools as well as parameters tested. Remember, a methods section should provide enough details such that the study can be reproduced.
    Results. Describe the results of your trials. Include figures and/or tables as appropriate.
    Discussion. Provide a synthesis of your results and how they fit in with previous studies. Include recommendations for when a particular tool should be used. Discuss limitations of your analysis.
    References. References cited in the text must be listed. References should be formatted using the PLOS Style (http://journals.plos.org/plosone/s/submission-guidelines#loc-references).

110 points are available. Your final grade will simply be the sum of all of the points you have earned over the course of the semester. The minimum points (pts) needed for each grade are listed in the table below. Points are rounded to the nearest tenth.

A ≥ 93 pts	B+ ≥ 87 pts	B- ≥ 80pts	C ≥ 73pts	D+ ≥ 67pts	F < 63pts
A- ≥ 90 pts	B ≥ 83 pts	C+ ≥ 77pts	C- ≥ 70pts	D ≥ 63pts	

Late/Missed Assignment Policy: Due dates and times for each assignment will be posted on ROSALIND or Sakai. LATE ASSIGNMENTS ARE TYPICALLY NOT ACCEPTED. **Extra points are built into the course for this reason**. Exceptions may be made at the discretion of the instructor.

Class Conduct: One important aspect of your education is learning to respect the rights of others. Please respect others by (1) allowing all classmates the right to voice their questions or opinions without fear of ridicule and (2) not making objectionable (e.g., gendered, racial, or ethnic) comments.

Diversity, Equity, and Inclusion Statement: Science strives to be objective. However, science is historically built on a small subset of privileged voices and thus includes both explicit and implicit biases. My goal is that students from all backgrounds and perspectives are well served by this course. The diversity students bring to bioinformatics should be viewed as a resource, strength and benefit as we pursue knowledge in the service of humanity. I intend to present materials and activities that are respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, religion or no religion, and culture. Your suggestions for improving inclusion are encouraged and appreciated.

Privacy Statement: Assuring privacy among faculty and students engaged in online instructional activities helps promote open and robust conversations and mitigates concerns that comments made within the context of the class will be shared beyond the classroom. As such, recordings of instructional activities occurring during synchronous Zoom sessions will be used solely for internal class purposes by the faculty member and students registered for the course, and only during the period in which the course is offered.

Intellectual Property: All lectures, videos, slides, and other instructional materials in this course are the intellectual property of the instructor. As a result, they may not be distributed or shared in any manner, either on paper or virtually without the instructor's written permission. Lectures may not be recorded without the instructor's written consent; all lectures will be recorded and available for students to view in Sakai through Zoom or Panopto. Recognizing that your work, too, is your intellectual property, the instructor will not share or distribute your work in any form without your written permission.

Statement of Intent: By remaining in this course, students agree to accept this syllabus as a contract and to abide by the guidelines outlined in the document.

Course Schedule

Week	Date	Topics/Readings	**DUE at start of class** (unless otherwise noted)
1	18-Jan	Algorithms and Python Review	
	20-Jan	Motif Detection Algorithms (Ch. 2)	
2	25-Jan	Work on Motif Problems	
	27-Jan	Code Check Assembly Algorithms (Ch. 3)	HW1: Motif Problems
3	1-Feb	Introduction to Linux/Class Server Discuss Project Expectations	
	3-Feb	Work on Assembly Problems	
4	8-Feb	Code Check Sequence Comparisons (Ch. 9)	HW2: Assembly Problems
	10-Feb	Journal Club – Dr. Wheeler	
5	15-Feb	Work on Sequence Comparison Problems	
	17-Feb	Code Check Clustering (Ch. 8)	HW3: Sequence Comparison Problems
6	22-Feb	Work on Clustering Problems	
	24-Feb	Code Check Hidden Markov Models (Ch. 10)	HW4: Clustering Problems
7	1-Mar	Work on HMM Problems	
	3-Mar	Code Check Project Choice Check	HW5: HMM Problems
8	8-Mar	***No class: Spring Break***	
	10-Mar	***No class: Spring Break***	
9	15-Mar	Journal Club, Work on Project	
	17-Mar	Journal Club, Work on Project	
10	22-Mar	Journal Club, Work on Project	
	24-Mar	Journal Club, Work on Project	
11	29-Mar	Journal Club, Work on Project	
	31-Mar	Journal Club, Work on Project	
12	5-Apr	Journal Club, Work on Project	
	7-Apr	Journal Club, Work on Project	
13	12-Apr	Journal Club, Work on Project	
	14-Apr	Journal Club, Work on Project	
14	19-Apr	Work on Project	
	21-Apr	Work on Project	
15	26-Apr	**Final Presentations**	
	28-Apr	**Final Presentations**	
Finals	6-May		Final Paper due 1:00PM

Schedule is subject to change at discretion of the instructor; changes will be published to Sakai.