Michigan Technological University

Course Syllabus

BL2700 – Principles of Bioinformatics
College of Science and Arts
Delivered In-person
Fall 2023

Instructor Information

Instructor:	Lindsay Putman, PhD, Research Assistant Professor
Office Location:	Dow 507
E-mail:	liputman@mtu.edu
Office Hours:	TBD or by appointment in Dow 507, or Zoom

Course Identification

Course Number:	BL2700
Course Name:	Principles of Computational Biology
Course Location:	MEEM 0202, Remote via Zoom on Fridays
Class Times:	MWF 1:00pm – 1:50pm

Course Description/Overview

Computational approaches pervade most disciplines. This course will cover the principles behind many computational biology including data gathering, handling, and statistics as well as these principles applied to ecological and molecular data (DNA and RNA sequences). Topics will cover guiding principles for data generation, curation, wrangling, and statistics as well as accessing biological data from databases, multiple sequence alignments and phylogenetic tree construction algorithms, tools for genome assembly and annotation, handling next-generation sequencing data.

Course Learning Objectives

The goal of Principles of Computational Biology (BL2700) is to provide foundational information and hands-on skills in commonly used tools in computational biology. There will be a particular focus on how biology, statistics and programming fit together to allow for in-depth understanding of biological systems. This will be accomplished through a combination of background lecture, in-class assignments, and projects. Through the course these activities will be designed to have students engage with and build tools for studying biological information. Assignments will also be designed to provide students with ability to hone their written and oral communication skills and evaluation of current literature. Upon successful completion of this course, students will be able to:

1. Explain and discuss the key principles behind best practices in data generation, handling of data, and statistical analysis of biological data.
2. Analyze the strengths and weaknesses of common bioinformatic algorithms.
3. Explain the interaction of biology, statistics, and computer algorithms in computational biology.
4. Interpret biological data through common software and custom programs.
5. Analyze current literature regarding the application of computational biology.

These learning objectives are also designed to link with core concepts and competencies in biology. In particular, these objectives are meant to provide a foundation in information flow, exchange, and storage.

Course Resources

Course Website(s)

- https://mtu.instructure.com/courses/1467420
- The Canvas website will be the home base for the online portion of this course. Here are some key places to look for information on the Canvas page
    - **Announcements:** Important reminders or course-wide information will be posted on the announcements part of the Canvas page
    - **Lectures:** This is where the PowerPoint slides for the lectures as well as the lecture recordings will be posted. Lectures will be posted later in the day following when the class happened. The lectures will have closed caption. To view closed caption, click the closed caption button at the bottom of the zoom window.
    - **Homework:** This is where the homework assignments will be posted.
    - **Modules:** This is where much of the content of the course will live. These will walk you through the content that we are covering in class as well as organize the practice and discussion sections of the class.
    - **Assignments:** Upcoming assignments will be posted on the Assignments tab of the page. This is the place to submit individual assignments.
    - **Discussions:** This is where participation and posts for discussion topics can be found.

Required Course Text

- Statistics for Ecologists Using R and Excel: Data Collection, Exploration, Analysis and Presentation (Data in the Wild): Mark Gardener
- Computing Skills for Biologists: A Toolbox: Stefano Allesina, Madlen Wilmes
- Additional reading will be posted on canvas site.

Grading Scheme

Grading System

Letter Grade	*Percentage*	*Grade points/credit*	*Rating*
A	93% & above	4.00	Excellent
AB	87% – 92%	3.50	Very good
B	82% – 86%	3.00	Good
BC	76% – 81%	2.50	Above average
C	70% – 75%	2.00	Average
CD	65% – 69%	1.50	Below average
D	60% - 64%	1.00	Inferior
F	59% and below	0.00	Failure
I	Incomplete; given only when a student is unable to complete a segment of the course because of circumstances beyond the student's control.		
X	Conditional, with no grade points per credit; given only when the student is at fault in failing to complete a minor segment of a course, but in the judgment of the instructor does not need to repeat the course. It must be made up by the close of the next semester or the grade becomes a failure (F). A (X) grade is included in the grade point average calculation as a (F) grade.		

Course Component	**Points**	**Percentage of Final Grade**
Exams (4 x 50 points)	200	45.5 %
Weekly Homework (12 x 10 points)	120	27.5 %
Genome Assembly project	50	11%
Participation (Attendance, Practice, and Discussions)	70	16%
Total Points	**440**	**100 %**

Exams (4 exams x 50 points each = 200 points)

Exams will be given through Canvas four times in the semester. The fourth exam will be given during the final exam period, but will only cover the material from the fourth unit. Each exam will be worth 50 points each. These exams will be a mix of multiple choice, short answers, and computer exercises to demonstrate competency with analysis. These are designed to primarily test knowledge of the theory behind the bioinformatics tools.

Weekly Homework (12 assignments x 10 points each = 120 points)

There will be weekly homework assignments to further develop skills related to the use of the bioinformatics tools. Data for the homework will be distributed through the canvas site. Homework

will either be in the form of quizzes or submitted as a word document through the assignments page on Canvas. Homework will be due midnight eastern time on Sunday.

Genome Assembly and Annotation Paper (50 points)

The end of the term project will be to assemble and annotate an unknown genome. Each student will be given raw data for a genome of an unpublished organism. You will be tasked with describing the quality of your assembly, key statistics about the genome's characteristics. You will identify the taxonomy of the organism of interest. Additional analysis will be performed such as metabolic modeling, gene ontology analysis or gene-specific analysis. This project will be written up into a paper of at least five pages (See posted example paper).

The paper will be in a standard font (Times New Roman, Arial. Cambria) size 12-point font. The paper should be double spaced. The paper should have the following sections (1) Introduction/Background which should describe the organism and information that is known about relatives of the organism. (2) Description of methods, which should include a description of each of the tools used including important parameters used. (3) Results, which includes details about the genomes statistics and comparison of different algorithms. (4) Conclusions, which should summarize the results from this project. It is essential that throughout the paper all information is properly cited.

Rubrics for these assignments will be posted on the Canvas page within the first month of class. The Genome project will be submitted in the form of a word document in the assignments page of the Canvas site.

Participation (70 points)

Each week there will be reading checks or practice labs that will be graded primarily for completion with feedback on the accuracy. The purpose of these are to provide opportunities for you to interact with the material and prepare for class.

Course Policies

Attendance: Participation in class during normal class time is expected. However, if you are not feeling well, please (1) Let Dr. Putman know, (2) Depending on the situation Dr. Putman may be able to provide you with a zoom link to log into class during the normal time. If neither 1 or 2 is possible, the lectures will be recorded and posted to the canvas site after class.

Make-up exams and Quizzes: Make-up exams and quizzes may be given only under certain extenuating circumstances.

Late Assignments: Points (10% of total) will be deducted per day for late submission of assignments. Assignments will not be accepted if they are more than 1 week late, unless the course instructor approves an excuse.

Academic Integrity Rules

Students may discuss homework assignments (if authorized), but are expected to individually work/write/solve any and all submitted work. All authorized resources used, including but not limited to internet sites (i.e. Chegg, Study Soup, Course Hero, etc.), should be appropriately cited. Please restrict all use of cell phones and/or other electronic devices during class to course-related

activities. The focus of class time should be interaction between students, and with the instructor. Any other unauthorized activities are likely to be distracting to other students and the instructor. Because it's important to everyone at Michigan Tech that academic standards be maintained, academic misconduct may result in an appropriate conduct sanction/educational condition(s) imposed by the Office of Academic and Community Conduct and/or in an academic penalty (lower grade/failing grade) imposed by the faculty.

For more details on academic integrity, please review the Academic Integrity Policy of Michigan Tech (http://www.admin.mtu.edu/usenate/policies/p109-1.htm).

University Policies

Student work products (exams, essays, projects, etc.) may be used for purposes of university, program, or course assessment. All work used for assessment purposes will not include any individual student identification.

Michigan Tech has standard policies on academic misconduct and complies with all federal and state laws and regulations regarding discrimination, including the Americans with Disabilities Act of 1990. For more information about reasonable accommodations or equal access to education or services at Michigan Tech, please call the Dean of Students Office at 906-487-2212 or go to http://www.mtu.edu/ctl/instructional-resources/syllabus/syllabus_policies.html.

Important Due Dates

Assignment	Date
Homework	Weekly
Exam 1	September 25th, 2023
Exam 2	October 18th, 2023
Exam 3	October 30th, 2023
Genome Paper	TBD – Finals week or week before
Exam 4	TBD - Finals Week

Tentative Course Schedule

Week	Class	Chapters (To be read prior to class)	Homework
Week 1			
8/28	*Introduction/ What is Computational Biology*		Homework 1 due
8/30	*Planning and Data Recording*	StatE: 1 & 2	
9/1	*Exploring data 1 – Beginning in R*	StatE: 3	
Week 2			
9/4	**No Class Labor Day**	CompB: 8	Homework 2 due
9/6	*Exploring data 2 – Beginning R 2*		
9/8	**No class K-day**		
Week 3			
9/11	*Exploring data 3 – Beginning R (tidyverse)*	CompB: 9A	Homework 3 due
9/13	*Exploring data 2 – Looking at distribution*	StatE: 4	
9/15	*Exploring data 3 – Graphing*	StatE: 6 and CompB 9B	
Week 4			
9/18	*Exploring data 4 – Which test is right*	StatE: 5	Homework 4 due
9/20	*Tests for differences 1 – Tests of difference*	StatE: 7	
9/22	*Tests for differences 3 – Correlations*	StatE: 8	
Week 5			
9/25	**Exam 1**		Homework 5 due
9/27	Databases 1	Web Reading	
9/29	*Introduction to R – 3* (Biostrings)	Web Reading	
Week 6			
10/2	*Algorithms for Multiple Sequence Alignments*	Web Reading	Homework 6 due
10/4	*Algorithms for Multiple Sequence Alignments*	Web Reading	
10/6	**Multiple Sequence Alignment - Lab**		
Week 7			
10/9	*BLAST 1*		Homework 7 due
10/11	*BLAST 2*		
10/13	*Molecular Evolution*		
Week 8			
10/16	*Molecular Phylogeny 1 (**not on exam 2**) – Guest or Pre-recorded lecture*		No HW
10/18	**Exam 2**		

10/20	*Molecular Phylogeny 2 – Neighbor joining*		
Week 9			
10/23	*Molecular Phylogeny 3 – Maximum Likelihood*		Homework 8 due
10/25	*Statistics applied to phylogenetic trees*		
10/27	**Lab 4: Molecular Phylogeny**		
Week 10			
10/30	**Exam 3**		No HW
11/1	*Genome Sequencing*		
11/3	**Lab 5 Unix Basics**		
Week 11			
11/6	*Algorithms in Genome Assembly 1*		Homework 9 due
11/8	*Algorithms in Genome Assembly 2*		
11/10	**Lab 6: Genome Assembly**	CompB:1	
Week 12			
11/13	*Genome Annotation 1 – BLAST*		Homework 10 due
11/15	*Genome Annotation 2 – Hidden Markov Models*		
11/17	**Lab 7: Genome Annotation**		
Week 13	**Thanksgiving Break**		
Week 14			
11/27	*Functional Predictions - KEGG*		Homework 11 due
11/29	*Microbiomes – 1*		
12/1	*Microbiomes – 2*		
Week 15			
12/4	*Microbiomes – 3*		Homework 12 due
12/6	*Microbiomes – 4*		
12/8	*Microbiomes – 5*		
Exam Week	**Exam 4 – Online, Date TBD based on your finals schedules**		