Office / contact info
CCDS 1536
bcleary@bu.edu

Course
CDS DS 526

Title
Critical Reading in Biological Data Science

Description
The goal of this course is to provide students with a framework, skills, and knowledge to critically evaluate research in biological data science. Biological research is rarely unequivocal in its findings; students will learn to systematically identify the claims advanced in research papers and evaluate whether each claim is established beyond a reasonable doubt by supporting evidence. We will examine papers that both meet and fail this test. In today's biology, to properly examine a paper in this way it is increasingly important to engage with the data provided as supporting evidence, and to critically examine the computational approach. Students will work with published data and computational tools. Further, students will learn to identify the ideology implicit in each paper, to understand how ideology shapes both the research questions and approach, and to imagine the same research under an alternative mindset. Classes will be split into lectures on background material for each paper and group discussions. Students will work in small groups to write a report on each paper. Each student will work on a final project to produce a critical review of a broader topic in the field.

Prerequisites
CDSDS 120, 121, and 122 or equivalent; ENGBE 562 or equivalent or experience with computational biology

Learning outcomes
A framework, skills, and knowledge to critically evaluate research in biological data science.

Assignments and grading structure
The course will consist of a deep dive, lasting several weeks, into each of seven papers.

For each paper students (together in small groups) will complete a critical evaluation worksheet to guide their evaluation. Completing the worksheet will involve identifying (i) basic information on claims and significance of the work, (ii) factual information presented in the study or relevant prior literature, (iii) a critique of the factual part of the study, (iv) interpretation of the factual findings (both the author's and the student's), (v) the arguments that move from findings or interpretation to claims, and (vi) an overall assessment.

In addition, for most of the papers we will devote one class session to a group hack of the data / model / code produced with each paper. The critical evaluation worksheet will be accompanied by results of this analysis.

The dive into each paper will conclude with an in-class, group discussion.

Finally, over the course of the second half of the term students will complete a final project consisting of a critical review of a broader subject in the field, to be selected in consultation with the instructor.

Final grades will be determined on the basis of worksheet evaluations (50%), final projects (25%), and in-class participation (25%).

Course policies

Since a large component of the class involves group discussion, in-person attendance is expected. In exceptional circumstances we will accommodate remote attendance with a hybrid class (in-person and zoom).

The BU Academic Code of Conduct is here: https://www.bu.edu/academics/policies/academic-conduct-code/. All students are required to familiarize themselves with this code, its definitions of misconduct, and its sanctions. Students should especially familiarize themselves with the section on plagiarism.

All written work in this course must be original to you. If you consult outside texts, or other forms of assistance, cite these sources in the proper format. This pertains to all external sources (books, journals, lectures, web sites, AI). We are required to report all suspected cases of plagiarism to the Academic Dean for review.

Academic integrity in computing coursework has some special aspects. Please review the examples of plagiarism as provided by the BU Computer Science department.

Logistics

Classes will frequently last 75 minutes (1:30 – 2:45), but we will occasionally take the full time (up to 3:15).

The group hack of the data / model / code will take place in-class on personal laptops. Work will be done in python, following along in a jupyter notebook.

Selection of Papers

- Topic 1: bacterial growth law (weeks 1-3)
    - Main paper
        - Cellular perception of growth rate and the mechanistic origin of bacterial growth law
    - Subjects
        - Quantitative biology, microbiology
- Topic 2: bacterial species interactions (weeks 3-5)
    - Main paper
        - Strength of species interactions determines biodiversity and stability in microbial communities
    - Subjects
        - Quantitative biology, microbiology, ecology
- Topic 3: neural circuits (weeks 5-7)
    - Main paper
        - The Code for Facial Identity in the Primate Brain
    - Subjects
        - Quantitative biology, computational biology, neuroscience

- Topic 4: brain development (weeks 7-9)
    - Main paper
        - Molecular logic of cellular diversification in the mouse cerebral cortex
    - Subjects
        - Quantitative biology, computational biology, genomics, neuroscience, developmental biology
- Topic 5: transcriptional kinetics (weeks 9-11)
    - Main paper
        - Mapping transcriptomic vector fields of single cells
    - Subjects
        - Quantitative biology, computational biology, genomics, immunology
- Topic 6: perturbation screens (weeks 11-13)
    - Main paper
        - D-SPIN constructs gene regulatory network models from multiplexed scRNA-seq data revealing organizing principles of cellular perturbation response
    - Subjects
        - Quantitative biology, computational biology, genomics, immunology
- Topic 7: networks (weeks 13-15)
    - Main paper
        - Network link prediction by global silencing of indirect correlations
    - Subjects
        - Quantitative biology, computational biology, genomics