Welcome to

CMSE 410/890 - Bioinformatics & Computational Biology

Arjun Krishnan

arjun@msu.edu | @compbiologist | thekrishnanlab.org

Week 1: Introductions

Course overview

- Introduction
- Course website
- Communication
- Course activities
- Tentative Schedule
- Course topics
- Wrap-up & what's due next

Land acknowledgement

Michigan State University occupies the ancestral, traditional and contemporary lands of the Anishinaabeg – Three Fires Confederacy of Ojibwe, Odawa and Potawatomi peoples. The university resides on land ceded in the 1819 Treaty of Saginaw.

Map: https://www.canr.msu.edu/nai/about/land-acknowledgements

Congratulations on surviving 2020 & 2021!

I'm amazed by and grateful for the seemingly limitless courage, hard work, and sacrifice of healthcare workers, other essential workers, and the people who stood up for social justice.

These two years have been incredibly tough for students, care givers (esp. parents of young children), and daily-wage workers — difficulties that were hugely compounded due to also belonging to systematically minoritized/disadvantaged groups.

In addition to these difficulties, my heart goes out to those
who also have suffered the loss of loved ones.

We need to continue doing our part including masking-up and vaccinating+boosting!

Introductions

- Please call me 'Arjun'. My pronouns are he/him/his.
- **arjun**@msu.edu | the**krishnan**lab.org | @**comp**biologist
- Assistant Professor
  - Dept. Computational Mathematics, Science, and Engineering
  - Dept. Biochemistry and Molecular Biology
- Research Interests: Computational genomics, Biomedical data science, Biological networks, Natural language analysis, Data integration, Machine learning
- Three things about me: 1) Always excited about learning new things, 2) Love hanging out with smart people like you, and 3) Enjoy chess, crosswords, sketching, music, and books.

Learn more about your fellow learners at https://bit.ly/compbio2021_flipgrid

What you should get out of this course

How to become a practicing computational biologist in 60+h?

- Introduction to the inner-workings of methods in bioinformatics and computational biology:
  - Analytical techniques, algorithms, and statistical/machine-learning approaches developed to address key questions in biology and medicine.

- At the end you should be able to:
  - Critically read bioinformatics / computational-biology literature.
  - Apply the methods you have learned to problems both within & outside biology.

Modules & Topics

1. Sequence alignment & pattern finding	• Dynamic programming; Substitution matrices; Fast Local Alignment
2. Genome assembly & annotation	• de Bruijin graphs; Hidden Markov models
3. Comparative genomics; Phylogenomics	• Suffix trees; Tree construction
4. Quantitative genetics	• Regularized linear regression; Statistical inference, Multiple testing
5. Regulatory genomics	• Gibbs sampling; Expectation-Maximization
6. Functional genomics	• Two-sample tests; Hypergeometric test; Unsupervised/supervised learning
7. Single-cell genomics	• Missing value imputation; Dimensionality reduction; Trajectory inference
8. Molecular dynamics; Structure prediction	• Molecular simulation; Maximum entropy modeling
9. Biological networks	• Measuring associations; Network inference; Graph theory, Label propagation
10. Modeling cellular pathways	• Wiring diagrams; Dynamical simulation, State Space, Bifurcation
11. Whole-cell models; Digital evolution	• Constraint-based modeling; Artificial life; Whole-cell models

Some general thoughts

Computational biology is extraordinarily inter-/multi-disciplinary.

- It is a melting pot of ideas from biology, computer science, physics, mathematics, and statistics.
- By definition, different subset of us will have a background in only some of these course topics and not others.
- And, by extension, you are all signing up to put yourself outside your technical comfort zone.

Some general thoughts

- **Conscious ignorance**: from unknown unknown → known unknown
  - Dunning-Kruger effect: knowing that something is unknown is as hard as knowing that thing!
  - The importance of feeling stupid: threshold of learning something new!

- **Intelligent persistence**
  - I don't understand this → What about this don't I understand?
  - Gaps in my knowledge → Gaps in collective knowledge

Prerequisites & Expectations

- CMSE 201 and two semesters of introductory biology (LB 144 and 145 OR BS 161 and 162 OR BS 181H and 182H, or equivalent).
  - Statistics at the level of STT 231 is strongly recommended.

- Basically, it would be assumed that you:
  - know how to code in one of the mainstream languages like Python or R,
  - have an understanding of basic statistics and probability, and
  - have studied basic genetics, molecular biology, and cellular biology.

Course website

bit.ly/compbio2022

- Home, Schedule, Lectures, and Assignments →
- Website and Communication
- Course Activities and Grading
- Philosophy, Presence, Conduct, Honesty, and Accommodations

- Calendar
- Links to:
  - Lecture slides
  - Learning materials
  - Assignments

NOT planning to use D2L!

Communication

compbio2022.slack.com

The primary mode of communication in this course (including major announcements) will be the course Slack account.

- #announcements
- #lectures
- #assignments
- #project

- #help
- #fun
- #health-workout
- #random

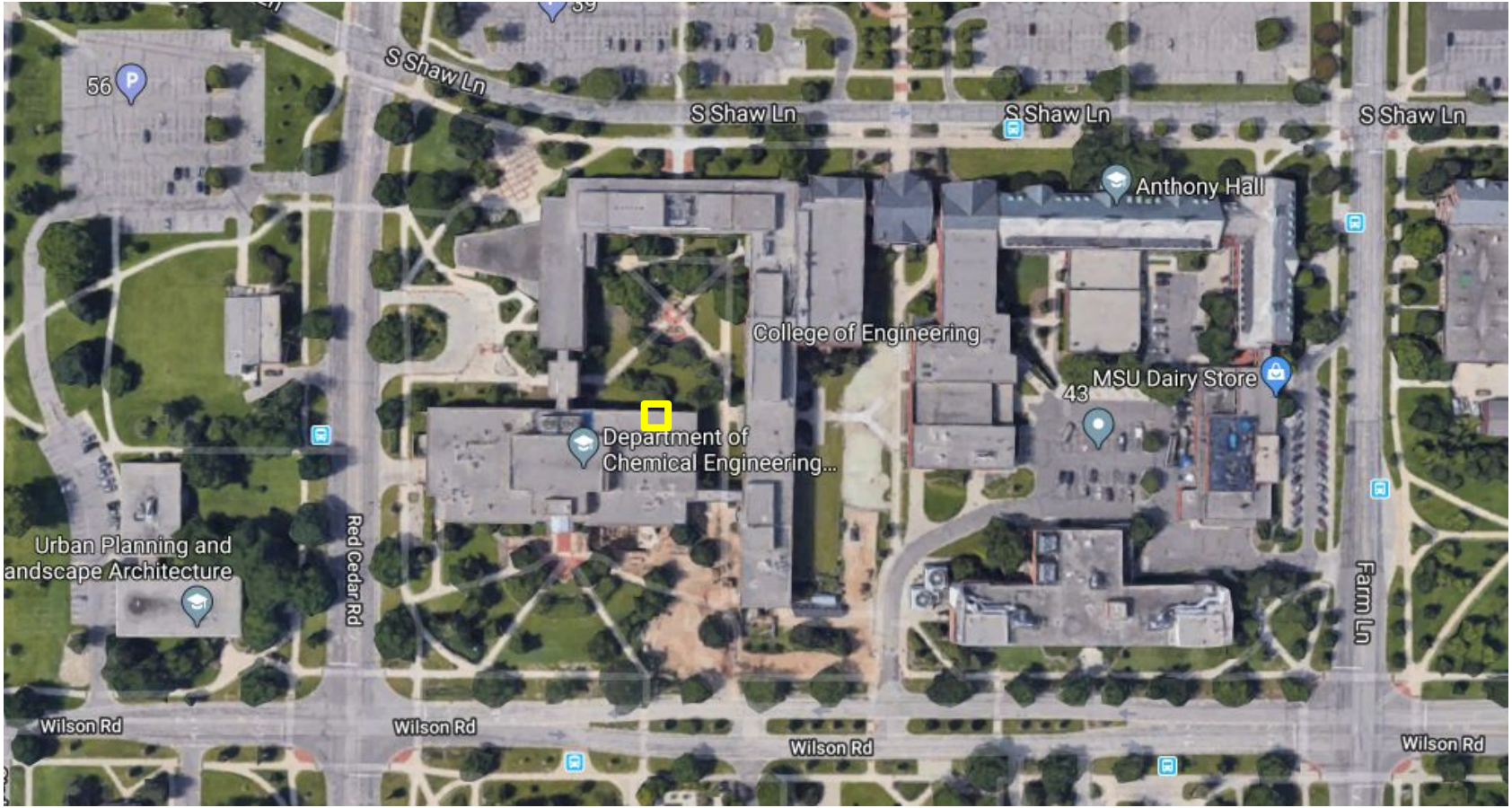
NOT planning to use D2L!

Offline hours

bit.ly/compbio2022_incoming

- Select convenient hours for offline discussion
  - Will give preference to enrolled students when picking the time.
  - Even if you're not able to make it to the designated hours, just messaging on Slack with your questions/concerns will work as well.

My office: 2507H Engineering Building (2nd floor)



Course activities

- Class presence | 30%
- Assignments | 35%
- Project | 35%

Class presence

- Completing pre-class readings / watching.
- Showing up to the (online) class on time and staying for the whole hour.
- Responding to prompts from Arjun that punctuate the lectures: Qs, polls, etc.
- Working in groups during in-class activities.
- Asking Qs about computational or biological concepts (because no one will have the perfect background).

Class presence

Stop me to ask questions. I love getting Qs from you!

I will extremely sad if I don't get Qs :(

In this class, **there's nothing wrong about being wrong**.

- Being wrong is an opportunity to learn something.
- So, when I pose questions, think about it and always take a shot.
- It not only helps you learn a new piece of information, but it also helps you calibrate how to think about that information, which is way more valuable.

Learning groups

A group of learners focused on a common goal and who respect & support each other in pursuit of that goal.

- Each group has a **dedicated slack channel**.
  - Post questions/comments; Help each other when needed; Coordinate meetings.
- **Meet** on Wednesday or Thursday to study/discuss that week's lectures & paper.
- Work on your assignments beforehand; Discuss w/ the group; **Submit individual assignments**.

Learning groups

Best practices:

- All class policies on **Presence, Conduct, & Honesty** apply to your shared space: https://github.com/krishnanlab/teaching/blob/master/2022-spring_compbio/policies.md.

- **Learn about your groupmates**: watch their FlipGrid profile & read their Research profile.
    - Learn their names & pronouns. [Please make your zoom/slack name to the name you go by.]

- **Prepare *before* your weekly meeting**:
    - Read the paper on your own once before discussing it when you meet.

Assignments

- **Lecture check-in**
  - After every Mon/Wed lecture, tell me three things you learned and anything that didn't make sense.
  - Mon → **Due** Noon, Tue | Wed → **Due** Noon, Thu
  - I will share a link to a google form.
- **Discussion check-in**
  - After every Fri paper discussion, tell me a few things about what you learned and anything that didn't make sense. Also tell me what did or did not work within your learning group and your discussion group.
  - **Due** 11:50p Fri
  - I will share a link to a google form.

Friday paper discussions

- I will provide you with a paper and guidelines on what to focus on.
- Learning groups meet and work together to understand the paper.
- Subset of members from each learning group form the discussion group for that week to lead the discussion of the paper based on a set of prompts/questions provided to you.

Typical week

Monday	Wednesday	Friday
In class: Arjun presents Lecture — Part 1.	**In class:** Arjun presents Lecture — Part 2.	**In class:** Arjun completes remaining materials (first 20 min). Students discuss paper of the week.
Outside class: Students submit lecture check-in by **Tue** Noon. Arjun posts details about Friday discussion.	**Outside class:** Students submit lecture check-in by **Thu** Noon. Learning groups meet & work together on **Wed/Thu**.	**Outside class:** Students submit discussion check-in by **Fri** 11:59p. Arjun posts (optional) Review quiz + Short video to review before next Monday.

Project

- A major goal of this course is to prepare your ability to perform original research in computational biology, and to present your ideas and research.

- Can be one of the following types:
  - Design and implement a new computational method for a task in biology.
  - Improve an existing method.
  - Perform an evaluation of several existing methods.
  - Develop a fully-reproducible documentation and codebase for an existing analysis in a paper.
    - If you are undergrad taking the 410 version of this course AND if you do not have any prior research experience, I strongly recommend this.

Project

1. Incoming survey | Mon, Jan 10
2. Research profile | Fri, Jan 21
3. Project topic | Fri, Feb 04
4. Project pre-proposal | Fri, Feb 11
5. Project proposal | Fri, Feb 18
6. Proposal reviews | Fri, Feb 25
7. Mid-term project presentation recordings | Wed, Mar 18
8. Mid-course Project presentation reviews | Fri, Mar 25
9. Final project report | Wed, Apr 29
10. Final project poster presentations | Tue, May 03, 12:45 – 2:45p

Teaching philosophy

Each class → an inclusive collaborative learning community.

- My goal is to make sure that our class is a space where all of you can:
  - Join in,
  - Breathe,
  - Be seen & heard,
  - Be curious, and
  - Openly engage with the ideas.

Teaching philosophy

I design and teach classes that:

- Maximize my learning,
- Help me identify gaps in my knowledge, and
- Find better ways of discussing each of the many complex/interesting ideas.

I will tell what parts of my understanding of these topics/ideas are works in progress and, hence, known-incomplete.

- I will try to be explicit about where the limits of my knowledge & understanding are.
- I have no problem saying "Hmm, I'm not sure. Let me think about this & get back to you" or "I have no clue now but, if you're interested, we can read a couple of sources together & revisit this."
- Correct me if/when I'm wrong.

Teaching philosophy

You absolutely belong in this community and you will be valued and respected.

My point of view is as follows:

- **Past**: Your unique background, training, and life experiences are your strengths that I and others can always learn from.
- **Present**: You have a life much bigger and multifaceted than your academic life within classrooms.
- **Future**: You are going to be my future colleagues within or outside academia.