MATH412/COMPSCI434/MATH713, Fall 2025 – Syllabus
Topological Data Analysis (Ling Zhou)

Instructor: Ling Zhou
- Email: ling.zhou@duke.edu
- Office: Physics 208

Class Times & Location: TuTh 3:05–4:20 PM, Gross Hall 318
Office Hours (tentative): TuTh 4:30–5:30 PM, Gross Hall 318 (or by appointment)

Course Materials:
1. The course textbook is ***Computational Topology for Data Analysis***, by Tamal Krishna Dey and Yusu Wang. Available for free online.
2. We will follow the structure of the following two courses closely:
    a. https://sites.google.com/ucsd.edu/dsc291-190-tda
    b. https://github.com/ZhengchaoW/DSC214-SPR24?tab=readme-ov-file
3. Useful resources for general knowledge of TDA:
    a. Computational Topology: An Introduction by Herbert Edelsbrunner and John Harer.
    b. A roadmap for the computation of persistent homology (2017), by Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod & Heather A Harrington
    c. Topological Data Analysis (2018), by Larry Wasserman

Prerequisites: Mathematics 221 (Linear Algebra and Applications) or equivalent.

Course Description:
This course introduces Topological Data Analysis (TDA), emphasizing both theory and applications. We will develop topological, geometric, and algebraic tools, beginning with persistence diagrams and their uses. Applications will appear early and often, alongside traditional theorem-proof mathematics and algorithmic development. The second half of the course explores more advanced topics in algebraic topology and their data applications, with exposure to relevant techniques from statistics and machine learning. Students will complete a final project applying TDA methods to a dataset of their choice (or a survey of methods).

Course Notes:
- Slides/notes will be posted on Canvas under Modules ⟶ Lecture Notes.
- The first half of the course relies heavily on the textbook, which students are expected to read outside of lecture.

Course Platforms and Learning Resources:
- **Canvas**: Course hub for announcements and lecture notes.
- **Gradescope**: Used for submitting homework, proposals, and final reports, as well as for grading and feedback; accessible through Canvas.

- **Ed discussion**: Accessible through Canvas. I will check Ed regularly and respond to questions at least once every weekday. Students are also encouraged to answer each other's questions when possible.
- Students are encouraged to work with each other. The [Academic Resource Center] (ARC) offers services to help you connect with peers: [Study Connect] and [Peer tutoring].

Grading Policy **(the percentages are adjusted on Sep 2nd)**:
There will be no exams; evaluation will be based on the following four parts:

- Homework 10.5% (individual work), due weekly in the first 7 weeks.
- Final project (group or individual work):
  - Proposal 10% (due 10/17)
  - Report 39.5% (due 12/4)
  - Presentation with Q&A 40% (due 12/4)

Use of AI Tools:
AI tools (e.g., ChatGPT, Copilot, etc.) may be used as a supplementary resource for brainstorming, checking code syntax, or clarifying background concepts. However:

- All submitted work must reflect **your own understanding**.
- Directly copying or lightly editing AI-generated solutions is not permitted and will be treated as academic dishonesty.
- You must be prepared to **explain any solution you submit**, whether on homework, in discussions, or during project presentations.
- For final projects, if AI tools meaningfully assisted your work (e.g., generating code, summarizing references), you should include a brief note in your report acknowledging how they were used.

Final Project:
The project can range from the applied (e.g., analyze a particular data set) to the theoretical (e.g., extend a particular theorem), and from the known (e.g,. review a published result) to the completely original. I will provide some examples/ideas/data sets for projects but, again, you have almost limitless freedom.

- Projects may be completed individually or in **groups of up to three**. Group projects are expected to demonstrate proportionally greater scope and depth of work.
- Each student (or group) is required to **meet briefly with me around Weeks 3-5** to discuss potential project topics.
- After a topic is chosen, you must submit a short **proposal (1-2 pages)** outlining your project plan.
- Upon completion of the project, you will submit a **final report (4+ pages)**, which should clearly describe the motivation of your project, what you have done, and your findings.
  - If your project takes the form of a survey, it needs to be **10+ pages**.
  - If your project involves code, you must also submit a **GitHub repository** containing your implementation, any test data (or links), and a short README with instructions for running the code.
- Each project includes a **5–10 minute in-class presentation followed by Q&A**. *Important: Part of your grade will be based on your ability to answer questions about both your project and related course concepts.*

- Presentations may also be recorded by the instructor for grading and reference purposes. Students who wish to create and share their own recordings (e.g., slides with voice-over) may do so, but this is optional.
- **Graduate students** are expected to submit longer documents: at least 2 pages for the proposal, 7+ pages for the final report, and 15+ pages for a survey project.

Here is the overleaf template for writing a proposal and a report
https://www.overleaf.com/read/bwdqbcxvdmtw
Page number requirements are specified with respect to the LaTeX default format (12pt font, 1-inch margins, single spacing).

Project ideas (added on Sep 2nd):
The goal of the class project is to encourage you to explore topological data analysis methods, whether through applications or theoretical research. Some possibilities

- use topological methods to analyze datasets of your interest
- theoretical improvements to existing results
- explore recent TDA-related papers in NeurIPS, ICML, or ICLR.

You can find plenty of papers on TDA applications in the Zotero database, which you may use as inspiration to improve upon an existing paper or apply to your dataset in a non-trivial manner. If you are having difficulty finding a suitable topic, don't hesitate to reach out to me and we can work together to find a topic that fits your background and interests.

Homework Assignments:
Homework will be assigned roughly weekly for the first half of the semester and due on **Thursdays at 11:59 pm**.

- Extensions & bonuses:
    - A 24-hour extension is granted upon request.
    - Submitting all assignments on time earns a +1% bonus to the homework average.
    - If a student typesets at least one homework assignment in LaTeX (excluding HW0), they will receive a +0.1% bonus applied to their overall homework grade of all homework assignments (adjusted on Sep 16th).
- Late policy: After any excused extensions, late submissions are penalized 5% per day.
- Submission: Upload completed work to Gradescope. Ensure pages are upright, legible, and properly **labeled/assigned** to each problem. Points may be deducted for submissions that are disorganized or difficult to read.
- Collaboration: You may discuss problems with peers, but all write-ups must be completed independently. Copying from peers or external sources is not allowed.
- Expectations:
    - Attempt problems before seeking help (office hours or Ed Discussion).
    - Clarity and logical completeness are required. Solutions must include full arguments, not just final answers.

Communication:

It is the student's responsibility to check their email daily (on weekdays) to get any important course updates, (for Outlook users, please check the "Other" folder, as announcements from Canvas sometimes get routed there). Log in periodically on Gradescope to make sure that exam and homework grades are recorded correctly; *any regrade requests should be submitted within one week of the time an assignment is returned.*

The instructor will plan to respond to emails within 2 business days. If a student has not received a response within that time frame, the student should send a follow-up email.

Attendance:

- Students are encouraged to attend all classes. A student absent from class bears full responsibility for all subject matter and procedural information discussed in class.
- Students who miss midterm exams or assignments due to a scheduled varsity athletic trip or religious holiday should submit an online NOVAP or RHoliday form, respectively, at least a week ahead of time and decide with the instructor how to make up the work.
- Those with a personal emergency or bereavement should inform their academic dean and instructor. Please contact the instructor as soon as possible after returning to schedule make-up work.
- Community Standard sanctions apply for abuse of this procedure.

Accommodations:

Students with disabilities who believe they may need accommodations in this class are encouraged to contact the Student Disability Access Office (SDAO) as soon as possible to better ensure that such accommodations can be implemented in a timely fashion. Once accommodations are approved, students should share their accommodation letter with me so that we can work together to ensure access to course materials and activities. For additional academic support and learning resources, students may also contact the Academic Resource Center (ARC): https://arc.duke.edu.

Academic Integrity:

Academic dishonesty on exams, plagiarism on homework and projects, copying homework, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty and will not be tolerated. They also violate Duke's Community Standard and will be referred to the Office of Student Conduct as described here in the Academic Integrity Council. Additionally, there may be penalties for their final course grade.

Tentative Schedule: see MATH412/COMPSCI434/MATH713 - Fall 2025 Schedule for the most up-to-date schedule.

Note: *the syllabus and course schedule are subject to change. Any changes to the syllabus and/or course schedule will be relayed to the students through Canvas and/or email.*

Wk	Dates	CTDA Sections	Topics Tues	Topics Thurs	HW due Thursdays	Notes
1	8/26, 8/28	Introduction, 1.1,	Introduction to the course	Topological spaces, maps, homeomorphisms		

		1.2, 1.3				
2	9/2, 9/4	1.4, 1.5, 2.1	homotopy, manifolds	simplicial complexes	HW0 due	
3	9/9, 9/11	2.2, 2.3, 2.4	cubical complexes, CW complexes, common choices	simplicial complexes: common choices, chains, cycles	HW1 due	Drop/Add ends 9/5. Meet with instructor to discuss project ideas.
4	9/16, 9/18	2.4, 2.5, 3.1	chains, cycles, boundaries, homology groups	matrix reduction algorithm, functoriality of homology	HW2 due	
5	9/23, 9/25	3.2, 3.3, 3.4	filtrations, persistent homology, barcodes, persistence module	persistent betti number, persistent matrix reduction algorithm	HW3 due	
6	9/30, 10/2	3.2, 4.1, 6.1	bottleneck distance, interleaving distance	Gromov-Hausdorff distance and stability	HW4 due	
7	10/7, 10/9	6.2, 6.3	point cloud data (PCD), homology inference	data sparsification for PCDs	HW5 due	Midsemester grades due 10/10.
8	10/16		Fall break	TDA applications in time series data	Proposal due 10/17	
9	10/21, 10/23	Chapter 7	TDA applications in Reeb graphs	TDA applications in graphs		
10	10/28, 10/30	Chapter 8	TDA applications in graphs	TDA applications in graphs		
11	11/4, 11/6	Chapter 9	Mapper	Mapper		
12	11/11, 11/13	Chapter 13	TDA and ML	TDA and ML		W deadline 11/11.
13	11/18, 11/20	Chapter 10	Discrete Morse theory	Discrete Morse theory		
14	11/25		TBD	Thanksgiving		
15	12/2, 12/4		Final presentation	Final presentation	Report due 12/4	