MIE524 Data Mining

Fall 2024

Course Syllabus

Instructor: Dr. Eldan Cohen, BA8134, ecohen@mie.utoronto.ca

Teaching Assistants:

- Riley Moher, riley.moher@mail.utoronto.ca

Times:

	Day	Time	Location
Lecture	Wednesday	1pm-2pm	WB 119
Lecture	Thursday	9am-11am	GB 221
Lab	Thursday	4pm-6pm	MB 123

Prerequisites:

- MIE350H1 or equivalent
- MIE236H1/ECE286H1/ECE302H1 or equivalent
- MIE245H1 or equivalent

Course Description:

Introduction to data mining and machine learning algorithms for very large datasets; Emphasis on creating scalable algorithms using MapReduce and Spark, as well as modern machine learning frameworks. Algorithms for high-dimensional data. Data mining and machine learning with large-scale graph data. Handling infinite data streams. Modern applications of scalable data mining and machine learning algorithms.

Textbook:

Mining of Massive Datasets 3rd Edition. Leskovec, Rajaraman, and Ullman. Cambridge University Press. 2020.

The book is available on http://www.mmds.org/

Grading:

- Assignments:
    - Deliverables: 20% of your final grade
    - Post-assignment in-class quizzes: 15% of your final grade
- Midterm: 25% of your final grade
- Final Exam: 40% of your final grade
    - **The Final Exam is mandatory and will result in course grade of incomplete (INC) assigned on the transcript if not attempted.**

Important Dates:

- **Assignment submission dates and in-class quizzes dates will be announced on Quercus.**
- **Midterm and Final exam dates will be announced on Quercus when they are finalized.**

Preliminary List of Topics:

The list of topics below is subject to change without notice.

- MapReduce and Spark
- Frequent itemsets and Association rules
- Neural networks and auto-encoders
- Large-scale supervised Machine Learning
- Locality Sensitive Hashing (LSH)
- Clustering
- Dimensionality reduction
- Analysis of massive graphs
- Link Analysis (PageRank)
- Recommender Systems
- Selected additional topics to be decided (e.g., Data Streams, Large Language Models)

Learning Outcomes:

1. **Objectives**. By the end of this course, students will be able to:

a. Understand the role of modern distributed computing and machine learning frameworks in creating scalable algorithms for large datasets.
b. Understand the algorithmic principles behind popular scalable data mining and machine learning approaches.
c. Be familiar with widely used, real-world applications of such approaches.
d. Become proficient in tools that support the development of scalable data mining and machine learning algorithms (MapReduce, Spark, and modern machine learning frameworks).
e. Become proficient in the methodology and tools for working with different types of data (high-dimensional, graph data, infinite data streams).

Course policies: re-grading, late submissions, petitions:

Re-grading: All requests for re-grading must (1) be submitted in writing (by email) to the instructor; (2) before the due date for re-grading requests (which is one week after the grades are posted). Only the instructor can make re-grading decisions, and re-grading requests will be given consideration only when the above conditions are satisfied.

Late submissions: Assignments submitted up to 48 hours late will be given a 30% late penalty (of the submission maximum mark). Projects submitted 48 hours late or more will be given a mark of zero.

Petitions: All petitions related to course term work should be submitted online through http://uoft.me/termworkpetition

Academic Integrity:

The University of Toronto expects you to be a full member of the academic community and to observe the rules and conventions of academic discourse. In particular, all the work you submit must be your own and no part of your submitted work should be prepared by someone else. Plagiarism or any other form of cheating in examinations, tests, assignments, or projects, is subject to serious academic penalty (e.g., suspension or expulsion from the faculty or university). Also, doing anything for the purpose of aiding or assisting another student to commit plagiarism is an offence which makes both parties liable for bearing the penalties and consequences. The full text of the policy that governs Academic Integrity at U of T (the Code of Behaviour on Academic Matters) can be found at:

www.governingcouncil.utoronto.ca/policies/behaveac.htm

For more information, please see the U of T Academic Integrity website.

Here are a couple of general guidelines to help you avoid plagiarism:

- Never look at another student's homework. Never show another student your solution. This applies to all drafts of a solution and to incomplete and even incorrect solutions.
- Keep discussions with other students focused on *concepts* and *examples*. Any code or solutions that you submit should be yours alone.
- Do not post any of your assignment questions in a private or public online discussion forum, social media and messaging groups, or website in order to solicit solutions from others.

University Land Acknowledgement

I wish to acknowledge this land on which the University of Toronto operates. For thousands of years, it has been the traditional land of the Huron-Wendat, the Seneca, and the Mississaugas of the Credit. Today, this meeting place is still the home to many Indigenous people from across Turtle Island and we are grateful to have the opportunity to work on this land.

Learn more about Canada's relationship with Indigenous Peoples here.

Indigenous Students' Supports

If you are an Indigenous engineering student, you are invited to join a private Discord channel to meet other Indigenous students, professors, and staff, chat about scholarships, awards, work opportunities, Indigenous-related events, and receive mentorship. Email Professor Bazylak if you are interested.

Indigenous students at U of T are also invited to visit Nations House's (FNH) Indigenous Student Services for culturally relevant programs and services. If you want more information on how to apply for Indigenous specific funding opportunities, cultural programs, traditional medicines, academic support, monthly social events or receive the weekly newsletter, go to the FNH website, email or follow FNH on social media: Facebook, Instagram, or TikTok. A full event calendar is on the CLNX platform. Check CLNX often to see what new events are added!

Wellness and Mental Health Support

Your personal wellness and mental health are important. The University of Toronto and the Faculty of Applied Science & Engineering offer a wide range of free and confidential services that can support your

well-being.

As a U of T Engineering student, you have a Departmental Undergraduate Advisor or a Departmental Graduate Administrator who can support you by advising on personal matters that impact your academics. Other resources that you may find helpful are listed on the U of T Engineering Mental Health & Wellness webpage, and a small selection are also included here:

- U of T Engineering's Student & Community Wellness Coordinator
- Health & Wellness and the On-Location Engineering Wellness Counsellor
- Health & Wellness Peer Support Program
- Accessibility Services & the On-Location Advisor
- Graduate Engineering Council of Students' Mental Wellness Commission
- SKULE™ Mental Wellness
- U of T Engineering's Learning Strategist and Centre for Learning Strategy Support
- Registrar's Office and Scholarships & Financial Aid Office & Advisor

We encourage you to access these resources as soon as you feel you need support; no issue is too small. You may reach out to the counsellors at U of T Telus Health Student Support for 24/7 free and confidential counselling support.

If you find yourself feeling distressed and in need of more immediate support visit uoft.me/feelingdistressed or U of T Engineering's Urgent Support – Talk to Someone Right Now.

Accommodations

The University of Toronto supports accommodations for students with diverse learning needs, which may be associated with mental health conditions, learning disabilities, autism spectrum, ADHD, mobility impairments, functional/fine motor impairments, concussion or head injury, visual impairments, chronic health conditions, addictions, D/deaf, deafened or hard of hearing, communication disorders and/or temporary disabilities, such as fractures and severe sprains, or recovery from an operation.

If you have a learning need requiring an accommodation the University of Toronto recommends that students register with Accessibility Services as soon as possible.

We know that many students may be hesitant to reach out to Accessibility Services for accommodations. The process of accommodation is private; we will not share details of your needs or condition with any instructor.

If you feel hesitant to register with us, we encourage you to reach out for further information and resources on how we can support. It may feel difficult to ask for help, but it can make all the difference during your time here.

Phone: 416-978-8060
Email: accessibility.services@utoronto.ca

Equity, Diversity and Inclusion

Looking for community? Feeling isolated? Not being understood or heard?

You are not alone. You can talk to anyone in the Faculty that you feel comfortable approaching, anytime – professors, instructors, teaching assistants, first-year or upper years academic advisors, student leaders or the Assistant Dean of Diversity, Inclusion and Professionalism.

You belong here. In this class, the participation and perspectives of everyone is invited and encouraged. The broad range of identities and the intersections of those identities are valued and create an inclusive team environment that will help you achieve academic success. You can read the evidence for this approach here.

You have rights. The University Code of Student Conduct and the Ontario Human Rights Code protect you against all forms of harassment or discrimination, including but not limited to acts of racism, sexism, Islamophobia, antisemitism, homophobia, transphobia, ableism, classism and ageism. Engineering denounces unprofessionalism or intolerance in language, actions or interactions, in person or online, on- or off-campus. Engineering takes these concerns extremely seriously and you can confidentially disclose directly to the Assistant Dean for help here.

Resource List:

- Engineering Equity, Diversity & Inclusion Groups, Initiatives & Student Resources
- Engineering Positive Space Resources
- Request a religious-based accommodation here
- Email Marisa Sterling, P.Eng, the Assistant Dean, Diversity, Inclusion & Professionalism here
- Make a confidential disclosure of harassment, discrimination or unprofessionalism here or email engineering@utoronto.ca or call 416.946.3986
- Email the Engineering Society Equity & Inclusivity Director here
- U of T Equity Offices & First Nations House Resources