

Artificial Intelligence for Drug Discovery

Important Links

- **Course web site:** https://www.cs.toronto.edu/~cmaddis/courses/csc2541_w26/
- **Quercus:** <https://q.utoronto.ca/courses/419279>

Course materials (schedule, slides, readings, assignments) can be found on the course web site.

Overview

Artificial intelligence is having a broad impact on society, and there is a push to harness these technologies for the purposes of drug discovery and development. This graduate course is a pilot course whose objective is take students through an exploration of the opportunities and challenges in AI for drug discovery and to provide a launchpad for genuinely inter-disciplinary, collaborative projects. It is delivered jointly through two sister courses: CSC2541 in Department of Computer Science and (PCL3107, PCL3108) in the Department of Pharmacology & Toxicology.

The scope of the course includes structural biology, phenomics, and genomics. With facilitation, students will conduct hands-on active learning assignments using simplified datasets to solidify their understanding of drug discovery data and problems. A course project, carried out in teams comprised of students from both departments, is the centrepiece of the course's work. The project will involve co-designing the research questions, research designs, and datasets of authentic projects at the intersection of AI and drug development. Tutorials will involve co-working sessions facilitated with expert guidance.

Students in the Computer Science Department will appreciate the priorities and research approaches of drug discovery scientists. Computer science students will share their domain-specific expertise and their perspectives about modern machine learning algorithms with their counterparts in Pharmacology & Toxicology. Students in the Pharmacology & Toxicology Department will appreciate the priorities and research approaches of AI scientists. Pharmacology and Toxicology students will share their domain-specific scientific expertise and their perspectives about experimental validation with AI scientists.

Prerequisites

This is a graduate course designed to explore the challenges in opportunities in AI for drug discovery. So, while there are no formal prerequisites, the course does assume a certain level of familiarity with key concepts, depending on the student's home department. Computer science students are expected to be proficient in machine learning, for example through having taken a previous course in machine learning such as CSC311 or STA314 or ECE421. Pharmacology & toxicology students are expected to have familiarity with biostatistics, ideally with large datasets. Students are not required to have prior experience with Machine Learning, but it is helpful to have experience/interest in upskilling in one or more of these areas: Foundations of Linear Algebra, Python, Introduction to Machine Learning.

Teaching Staff

CSC Instructor (giving PCL lectures): Chris J. Maddison

PCL Instructors (giving CSC lectures): Rebecca Laposa and Jean Martin Beaulieu

Teaching Assistants: Micaele Elisa Consens (CSC2541), Ella Miray Rajaonson (CSC2541), Stef Vislavski (PCL3107,PCL3108)

Staff e-mail: ai4dd@cs.toronto.edu

Please send course-related e-mails to the instructor/staff emails above. If you have a private matter that you would like to discuss with the instructor, CS students can email cmaddis@cs.toronto.edu. PCL students can email rebecca.laposa@utoronto.ca.

Schedule

Lecture Times

- PCL3107H: Wednesdays 1:00PM-3:00PM.
- CSC2541H, PCL3108H: Thursdays 1:00PM-3:00PM.
- Joint Tutorials for all students: Fridays 3:00PM-4:00PM.

Lecture and Tutorial Structure The first 6 weeks of the course are divided into three major sections that address three major components of artificial intelligence in drug discovery. See the course web page for information about topics. Each of the three sections is addressed in lectures and has an associated Module (Google Colab) to reinforce and explore the concepts further. The modules have a pre-module section, the main module to be completed in interdisciplinary teams during tutorial time, and a post-module to be completed individually. Attendance and active engagement at each and every tutorial is expected of all students in order to fulfil the learning objective of interdisciplinary collaboration.

A course project, carried out in teams comprised of students from both departments, is the centrepiece of the course's work. The project will involve co-designing the research questions, research designs, and datasets of authentic projects at the intersection of AI and drug development. Tutorials will involve team co-working sessions facilitated with expert guidance (see Schedule). Student teams will submit their letter of intent research proposal in the first 6 weeks and conduct the research project, including milestone writeups, in the second 6 weeks.

We recommend that you read at least the core readings and as many of the additional papers as possible. We won't check whether you've read the assigned readings, but you will get more out of the course if you do.

Recordings

Lecture recordings will be generated and posted automatically on the OCCS Student App.

This course, including your participation, will be recorded on video and will be available to students in the course for viewing remotely and after each session.

Course videos and materials belong to your instructor, the University, and/or other sources depending on the specific facts of each situation, and are protected by copyright. Do not download,

copy, or share any course or student materials or videos without the explicit permission of the instructor.

For questions about recording and use of videos in which you appear please contact your instructor.

Course Evaluation—CSC2541H

- 22.5% – Modules/code notebooks (3 X 7.5%, due Jan 23, Feb 6 and Feb 20)
- 20% – Team Project proposal (Letter of Intent) (due Feb 23)
- 22.5% – Team Project milestone writeups (2 X 11.25%, due Mar 9 and 30)
- 20% – Team Project Oral Presentation (presentations Apr 1-3)
- 15% – Participation (ongoing)

Details will be posted on the course web site.

Course Evaluation—PCL3107H

- 45% – Modules/code notebooks (3 X 15%, due Jan 19, Feb 2 and Feb 16)
- 40% – Team Project proposal (Letter of Intent) (due Feb 23)
- 15% – Participation (ongoing)

Details will be posted on the course web site.

Course Evaluation—PCL3108H

- 45% – Team Project milestone writeups (2 X 22.5%, due Mar 9 and 30)
- 40% – Team Project Oral Presentation (presentations Apr 1-3)
- 15% – Participation (ongoing)

Details will be posted on the course web site.

Submission Policies

There are 3 individual assignments to submit (Modules) as well as team assignments: a project proposal, two project milestone reports and a final project oral presentation.

Format. The presentation slides, project proposal, and final project write-up must be submitted in PDF format through Quercus. We encourage typesetting using L^AT_EX, but other formats are acceptable as long as they are legible.

Lateness. Assignments will be accepted up to 3 days late, but 10% will be deducted for each day late, rounded up to the nearest day. No credit will be given for assignments submitted after 3 days. Extensions will be granted only in special situations, and you will need a Student Medical Certificate or a written request approved by the instructor at least one week before the due date.

Collaboration policy. Collaboration on the research proposal and final project is essential and interdisciplinary teams should be finalized during the week of Jan 19 or before. The first tutorial social mixer is designed to give students a chance to meet each other and form teams. Each team needs to have member(s) from the Department of Computer Science and the Department of Pharmacology and Toxicology. Collaboration during the tutorial Modules (Google Colabs) is

highly encouraged. The presentation for the final project should list the contributions of each team member.

Remarks. Remark requests will be considered by the same TA who marked the assignment. The deadline for requesting a remark is one week after the marked assignments are returned. Remarks may result in a decrease in the grade.

Policy on The Use of Artificial Intelligence

Large-language-model-based chatbots like ChatGPT and Claude can be very useful educationally, and you're encouraged to take advantage of them. The only restriction is that **you are not allowed to use them in a way that trivializes an assignment**, such as asking them how to solve a problem.

Along with each assignment, **you must submit (on Quercus) any chat transcripts directly related to the assignment**. We'll err on the side of permissiveness (i.e. you won't be penalized as long as you're acting in good faith), but we may revise the policy if we see that use of chatbots is reducing the educational value of the assignments. If a chatbot substantially more powerful than the current frontier models is released during the term, we may amend the policy to address that.

Computing Support

As a graduate-level course, we are assuming that many students in this class have a laboratory affiliation which grants them access to computing resources which they may use for the purposes of this class. If you do not have access to compute via a laboratory affiliation, please consider using Google Colab (<https://research.google.com/colaboratory/>) to run your experiments, or contact the teaching staff **as early as possible** and we will do our best to provide some for you.

Accessibility Support

We are committed to making our classroom an accessible environment. If you require additional academic accommodations, please contact UofT Accessibility Services as soon as possible, studentlife.utoronto.ca/as, and contact the course staff.

Auditing Policy

We are not allowing auditors in this course.