CSC311 - Introduction to Machine Learning (Winter 2023)

Overview

Machine learning (ML) is a set of techniques that allow computers to learn from data and experience rather than requiring humans to specify the desired behaviour by hand. ML has become increasingly central both in AI as an academic field and industry. This course provides a broad introduction to some of the most commonly used ML algorithms. It also introduces vital algorithmic principles that will serve as a foundation for more advanced courses, such as CSC412/2506 (Probabilistic Learning and Reasoning) and CSC413/2516 (Neural Networks and Deep Learning).

We start with nearest neighbours, the canonical non-parametric model. We then turn to parametric models: linear regression, logistic regression, soft max regression, and neural networks. We then move on to unsupervised learning, focusing in particular on probabilistic models, but also principal components analysis and K-means. Finally, we cover the basics of reinforcement learning.

Announcements

Welcome to CSC311! A math diagnostic will be released in the next few days. In the meantime, sign up for Piazza. This website will have minor changes throughout the week.

Course Sections

There are two sections of this course being offered this term. As of January 2023, we plan to have in-person lectures, tutorials, office hours, midterm and final exam. This may, given the COVID-19 situation, be subject to change by the university.

Please attend your assigned lecture section. We strongly encourage students to attend the tutorials although they are optional. Auditing is not allowed this term without express written permission by the instructor.

Section	Instructor	Lecture	Tutorial
0101	Michael Zhang	M 11:00 - 13:00, BA1190	W 11:00 - 12:00, BA1190
0102	Chandra Gummaluru	M 14:00 - 16:00, BA1160	W 14:00 - 15:00, BA1160

A Note on COVID-19

Although the pandemic has diminished somewhat, all indications are that we are not yet out of the woods. The university no longer requires the use of masks on its premises but encourages it where it is impossible to maintain physical distancing, such as in classrooms and offices.

We strongly recommend that you continue to **wear masks** during lectures, tutorials, and office hours out of consideration for the health of others. We also strongly encourage you to **get vaccine booster shots** whenever possible. The instructors plan to wear masks when in close proximity with students, such as when answering questions after lectures or during office hours. However, we may take off our masks when lecturing if we are at a safe distance from students.

Prerequisites

This course has the following pre-requisites:

- **Programming Basics**: CSC207/ APS105/ APS106/ ESC180/ CSC180
- **Multi-Variable Calculus**: MAT235/ MAT237/ MAT257 / (77%+ in MAT135, MAT136)/ (77%+ in MAT137)/ (67%+ in MAT157)/ MAT291/ MAT294 / (77%+ in MAT186, MAT187)/ (73%+ in MAT194, MAT195) / (73%+ in ESC194, ESC195)
- **Programming Basics**: CSC207/ APS105/ APS106/ ESC180/ CSC180
- **Linear Algebra**: MAT221/ MAT223/ MAT240/ MAT185/ MAT188
- **Probability**: STA237/ STA247/ STA255/ STA257/ STA286/ CHE223/ CME263/ MIE231/ MIE236/ MSE238/ ECE286

Grading Scheme

We will use the following grading scheme for the course.

Evaluation	Weight
Math Diagnostic Quiz	2%
Assignments	3 $\times$ 11% = 33%
Mid-Term Exam	20%
Ethics Module	5%
Final Exam	30%

Note that you must obtain a grade of **at least 40%** on the final exam to pass the course. For the Ethics Module, the 5% weight will be split into mini-assignments, as follows:

Mini-Evaluation	Weight	Marking
Pre-Module Survey	1%	Completion
Class Participation	0.5%	Automatically Given
Reflection on Class Activity	2%	Good-Faith Effort
Post-Module Survey	1.5%	Completion
In short, we expect almost all students to get the full 5%.		

Communication

There are many ways to get in touch with us.

- **Piazza**: https://piazza.com/utoronto.ca/winter2023/csc311 Fastest way to get answers to course related questions.
- **Course Email Address**: csc311-2023-01@cs.toronto.edu
- **Instructor Office Hours**:
    - Michael: 10-11am Wednesday, location TBA
    - Chandra: TBA
- **TA Office Hours**: These will be held for each assignment, the mid-term exam, and final exam. Please monitor the respective sections for dates/times.

Please follow these rules when you contact us:

- If your question is **course related and doesn't give away answers**, please post on Piazza publicly so the entire class can benefit from the answer.
- If your question is **course related and may give away answers**, please post on Piazza privately.
- For **remark requests**, please submit on MarkUs (for assignments) or contact us via the course email.
- For **special considerations requests**, please contact us via the course email.

TAs will hold office hours to help with assignments and project, as well as preparing for the midterm and the final exam. TA office hours will be posted in the respective sections.

Assignments

There will be 3 assignments in this course, posted below. Assignments will be due at 11:59pm on Wednesdays and submitted through MarkUs.

Computational Resources: We will use Python 3, and libraries such as NumPy, SciPy, and scikit-learn. You have two options:

1. The easiest option is probably to install everything yourself on your own machine.
    1. If you don't already have Python 3, install it.

       We recommend some version of Anaconda (Miniconda, a nice lightweight conda, is probably your best bet). You can also install Python directly if you know how.
    2. Optionally, create a virtual environment for this class and step into it. If you have a conda distribution run the following commands:

       ```
       conda create --name csc311
       source activate csc311
       ```
    3. Use `pip` to install the required packages

       ```
       pip install scipy numpy autograd matplotlib jupyter sklearn
       ```
2. All the required packages are already installed on the Teaching Labs machines.

Late Submission Policy: Everyone will receive 3 grace days, which can be used at any point during the semester. No credit will be given for assignments submitted after 3 days.

Collaboration Policy: Collaboration on assignments is not allowed. Each student is responsible for his/her own work. Discussion of assignments should be limited to clarification of the handout itself, and should not involve any sharing of pseudocode or code or simulation results. Violation of this policy is grounds for a semester grade of F, in accordance with university regulations.

Remark Policy: If you discover a marking error on an assignment, you can submit a remark request. We will consider remark requests **up to two weeks** after we release the marks for an assignment or the midterm. Please submit your remark request via MarkUs.

Mid-Term Examination

About half-way through the term, we will have a mid-term examination. You will be allowed to bring an aid-sheet (one-side 8.5" by 11"). More details will be posted as we get closer to the mid-term.

Final Examination

At the end of the term, we will have a FAS proctored final examination. You will be allowed to bring an aid-sheet (two-sides 8.5" by 11"). More details will be posted as we get closer to the final exam.

Course Project

For your final project, you will consider an application of the material we have learned in this course. You will work in groups of 2-3. The details will be posted later in the term.

Lecture and Tutorial Materials

You can find all of the relevant lecture/tutorial materials below. The suggested readings are optional. We have provided them in case you need alternate resources to understand the material. All of the textbooks listed below are freely available online.

- Bishop: Pattern Recognition and Machine Learning.
- Hastie, Tibshirani, and Friedman: The Elements of Statistical Learning.
- MacKay: Information Theory, Inference, and Learning Algorithms.
- Barber: Bayesian Reasoning and Machine Learning.
- Sutton and Barto: Reinforcement Learning: An Introduction.
- Deisenroth, Faisal, and Ong: Math for ML.
- Shalev-Shwartz and Ben-David: Understanding Machine Learning: From Theory to Algorithms.
- Murphy: Machine Learning: a Probabilistic Perspective.

Week	Topics Covered	Materials	Suggested Readings
1	**Lecture:** Nearest Neighbours **Tutorial:** Probability Review		

Academic Integrity

Academic integrity is essential to the pursuit of learning and scholarship in a university, and to ensuring that a degree from the University of Toronto is a strong signal of each student's individual academic achievement. As a result, the University treats cases of cheating and plagiarism very seriously. The University of Toronto's Code of Behaviour on Academic Matters outlines the behaviours that constitute academic dishonesty and the processes for addressing academic offences.

All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour on Academic Matters. If students have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, they are expected to seek out additional information on academic integrity from their instructors or from other institutional resources.

Special Considerations Policy

If you are unable to complete a course requirement due to extraordinary circumstances beyond your control, please apply for a Special Consideration by filling out this special considerations form and sending it to the course email with your supporting documentation. A special consideration request, particularly if it is not your first request in the course, would not be granted automatically.

Legitimate reasons to apply for a special consideration request:

- Late course enrollment
- Medical conditions (i.e., physical/mental health, hospitalizations, injury, accidents)
- Non-medical conditions (i.e., family/personal emergency)

A heavy course load, multiple assignments/tests scheduling during the same period, and time management issues are not appropriate reasons to grant special considerations. Such accommodations are meant for exceptional circumstances only and not as a means to catch up on term work. If you are having difficulty with stress and time management, please contact your college registrars, who can in turn suggest wellness counselling, academic advising, and/or learning strategists services.

Our special considerations policies are as follows.

- If you miss the midterm, we will shift the weight of the midterm to the final exam.
- If you miss an assignment deadline, we will shift the weight of the assignment to future assignments or to the final exam.
- If you are registered with accessibility services, your letter of accommodation will allow for an extension of up to 7 full days. However, due to the incremental nature of CS courses, granting such a long extension from the onset may cause you to fall behind and be at a disadvantage. As such, we will start by suggesting an initial 3-day extension. We will grant the 7-day extension later if necessary.