CSC311 - Introduction to Machine Learning (Summer 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. We will develop a mathematical foundation to understand and implement these algorithms.

Jump to:

Assignments
Lectures

Course Sections

There is only one section of this course being offered this term. We plan to have in-person lectures, tutorials, office hours, midterm and final exam. Occasionaly, we may make the lectures or office hours online.

Auditing is not allowed this term without express written permission by the instructor.

Section	Instructor	Lecture	Tutorial
101	Sayyed Nezhadi	Thu 18:00 - 20:00, BA1170	Thu 20:00 - 21:00, BA1170

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
Assignments (3)	35%
Project	15%
Midterm Exam	20%
Final Exam	30%

Note that you must obtain a grade of **at least 40%** on the final exam to pass the course.

Communication

There are many ways to get in touch with us.

- **Piazza**: https://piazza.com/utoronto.ca/summer2023/csc311h5 Fastest way to get answers to course related questions.
- **Course Email Address**: csc311-2023-05@cs.toronto.edu
- **Instructor Office Hours**:

    - 5-6pm Thursday, TBD
    - I will prioritize conceptual questions during instructor office hours.

- **TA Office Hours**: These will be held for each assignment, the midterm exam, and final exam. Please monitor the announcements 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).
- For **special considerations requests**, please contact us via the course email. There is a ticketing system which instructional support helps us process.

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 11:59 pm on Mondays or Fridays and submitted through MarkUs.

#	Materials	TA Office Hours
Assignment 1	To be relaesed.	TBD
Assignment 2	To be relaesed.	TBD
Assignment 3	To be relaesed.	TBD

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 scikit-learn
      ```

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 on the three assignments. No credit will be given for assignments submitted after 3 days. We will have a separate policy on the final project.

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.

Midterm Examination

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

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

To be relaesed

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 (ESL).
- 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:** Introduction / Nearest Neighbours	**Lecture:** Slides	Math for ML (skim) ESL: 1, 2.1-2.3, 2.5
2	**Lecture:** Decision Trees, Bias-Variance Decomposition **Tutorial:** Probability Review	**Lecture:** Slides	Bishop: 3.2 ESL: 2.9, 9.2 Course notes: Notes on Generalization, Bias-Variance Decomposition
3	**Lecture:** Linear Models 1 (Linear Regression) **Tutorial:** Linear Algebra Review	N/A	Bishop: 3.1 ESL: 3.1 - 3.2 Course notes: Linear Regression, Calculus
4	**Lecture:** Linear Models 2 (Logistic Regression) **Tutorial:** Bias-Variance Decomposition	N/A	Bishop: 4.1, 4.3 ESL: 4.1-4.2, 4.4, 11 Why momentum really works Course notes: Linear Classifiers, Training a Classifier
5	**Lecture:** Linear Models 3 / Review **Tutorial:** Optimization	N/A	Bishop: 5.1-5.3 Course notes: Multilayer Perceptrons
6	**Lecture:** Neural Networks **Tutorial:** Midterm Review	N/A	Neural Net Playground
7	**Midterm (Date TBA)**		
8	**Lecture:** Neural Networks 2 **Tutorial:** Pytorch	N/A	Backpropagation Convolutional neural networks ESL: 2.6.3, 6.6.3, 4.3.0 MacKay: 21, 23, 24 Probabilistic models
9	**Lecture:** Probabilistic Models **Lecture 2:** Deep Learning - Special Topics	N/A	TBA
10	**Lecture:** Multivariate Gaussians, GDA **Tutorial:** Linear Algebra Review Part 2	N/A	Math for ML Bishop: 12.1
11	**Lecture:** PCA, Matrix Completion	N/A	ESL: 14.5.1

Week	Topics Covered	Materials	Suggested Readings
	Tutorial: PCA		
12	**Lecture:** K-means, RL **Tutorial:** Final Review	**N/A**	MacKay: 20 Bishop: 9 Barber: 20.1-20.3 See slides for RL recs.

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.

Student Support Resources

- CSC Wellness Resources
- Student Mental Health Resources
- Accessibility Services
- Student Life
- Recognized Study Groups

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.