CS 480/680:
Introduction to Machine Learning
Fall 2024

Time and Location
480/680 Sec. 001: TR 1:00 - 2:20 PM MC 4021
480/680 Sec. 002: TR 8:30 - 9:50 AM MC 2017

Links
Homepage: github.com/kpc-simone/cs480-f24
Submissions (LEARN/CrowdMark): learn.uwaterloo.ca/d2l/home/1046818
Discussions (Piazza): piazza.com/uwaterloo.ca/fall2024/cs480680

Syllabus

Instructional Team

Instructor: Kathryn Simone
Office: DC 2126
Office Hours: See **Policies**.
Email: kpsimone@uwaterloo.ca

T.A.	Carter Blair	Matina Mahdizadeh Sani	Saber Malekmohammadi	Evelien Riddell
Email (@uwaterloo.ca)	cblair	m3mahdiz	s3malekm	eeboerst

Course Information

Course Description

Introduction to modeling and algorithmic techniques for machines to learn concepts from data. Generalization: underfitting, overfitting, cross-validation. Tasks: classification, regression, clustering. Optimization-based learning: loss minimization. regularization. Statistical learning: maximum likelihood, Bayesian learning. Algorithms: nearest neighbour, (generalized) linear regression, mixtures of Gaussians, Gaussian processes, kernel methods, support vector machines, deep learning, sequence learning, ensemble techniques. Large scale learning: distributed learning and stream learning. Applications: Natural language processing, computer vision, data mining, human computer interaction, information retrieval. [Note: Lab is not scheduled and students are expected to find time in open hours to complete their work.]

Course Goals and Learning Outcomes

By the end of this course students should be able to:

- Recognize and formulate a task as a ML problem;
- Identify and recommend suitable algorithms to tackle different ML problems;
- Implement foundational ML algorithms;
- Apply and evaluate ML algorithms on real datasets;
- Justify and critique choices in terms of ML principles;
- Describe ethical and safety issues of ML on society.

Course Resources and Communication

Below are links to the key resources you will need for this class. Please make sure to familiarize yourself with these platforms as they will be essential for your success in the course.

Course Homepage: All course materials are available on our GitHub repository: GitHub: CS480/680-F24
Please refer to this link regularly for the latest updates on lecture slides, video lectures, and reading materials.

Assignment Submissions: All assignments must be submitted through Waterloo LEARN (D2L). Ensure that you upload your work by the specified deadlines. Access the course on LEARN here: LEARN: CS480/680-F24

Questions, Discussions, and Announcements: For any questions, discussions, or announcements, we will be using Piazza. It is a great platform for interacting with your peers and getting guidance from the teaching team. Join the course on Piazza: Piazza: CS480/680-F24

Table 1: Course Outline (Tentative)

The table below lists the plan for the topics that will be covered in the course and the order in which they will be covered. Note that this list is **tentative**, and the order may be changed.

Lecture	Date	Topics	Assignments
0	05/09/2024	Introduction + Administrative Remarks	
1	10/09/2024	Halfspaces the Perceptron Algorithm	Assignment #1: Posted (Due: Sept 27th)
2	12/09/2024	Linear Regression and Convexity	
3	17/09/2024	Maximum Likelihood Estimation	
4	19/09/2024	k-means Clustering	
5	24/09/2024	k-NN Classification and Logistic Regression	
6	26/09/2024	Hard-margin SVM	Assignment #2: Posted (Due: October 14th)
7	01/10/2024	Soft-margin SVM	
8	03/10/2024	Kernel methods	
9	08/10/2024	Decision Trees	
10	10/10/2024	Bagging and Boosting	
	15/10/2024	NO LECTURE - MIDTERM BREAK	
	17/10/2024	NO LECTURE- MIDTERM BREAK	
11	22/10/2024	Expectation Maximization Algorithm	
12	24/10/2024	MLPs and Fully-Connected NNs	Assignment #3 Posted (Due November 8th)
	29/10/2024	NO LECTURE - MIDTERM EXAM	
13	31/10/2024	Convolutional Neural Networks	
14	05/11/2024	Recurrent Neural Networks	
15	07/11/2024	Attention and Transformers	
16	12/11/2024	Graph Neural Networks (Time permitting)	
17	14/11/2024	VAEs and GANs	Assignment #4 Posted (Due November 24th)
18	19/11/2024	Flows	Title: “Modern Topics and Trends”
19	21/11/2024	Contrastive Learning (Time permitting)	
20	26/11/2024	Robustness	
21	28/11/2024	Privacy (Saber Malekmohammadi)	
22	03/12/2024	Fairness	

Reference Materials

Suggested References

There is no required text for the course. Suggested sections from textbooks and/or other resources will be referred to on the course homepage prior to the lecture in question. Readings will be drawn most often from the following reference texts.

- An Introduction to Statistical Learning (ISL), by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani.
- Elements of Statistical Learning (ESL), by Trevor Hastie, Robert Tibshirani and Jerome Friedman.
- Understanding Machine Learning: From Theory to Algorithms (UML), by Shai Shalev-Shwartz and Shai Ben-David.
- Deep Learning (DL), by Ian Goodfellow, Yoshua Bengio and Aaron Courville.
- Dive into Deep Learning (D2L), by Aston Zhang, Zack C. Lipton, Mu Li and Alex J. Smola.

The following texts are excellent supplementary resources:

Mathematics for Machine Learning (MML), by Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong.

Course Requirements and Assessments

The course will have assessments in the form of assignments and exams, and for CS 680 only, there will be a final project.

Table 2: Grading Scheme

Assessment	TA (if applicable)	Assessment Date	Weighting (CS480)	Weighting (CS680)
Assignment 1	Saber Malekmohammadi	September 27	7.5%	7.5%
Assignment 2	Matina	October 14	7.5%	7.5%
Assignment 3	Carter	November 8	7.5%	7.5%
Assignment 4	Carter	November 22	7.5%	7.5%
Exams				
Midterm	All	October 29	30%	15%
Final	All	TBD	40%	30%
Project (CS 680 only)	Instructor			
Pitch	Evelien	September 19	N/A	2%
Proposal	Evelien	October 8	N/A	8%
Report	Instructor	December 3	N/A	15%
Total			100%	100%

Assignments

There will be 4 assignments, worth a total of 30% of the final grade. The key aim of assignments is to deepen understanding of a major subtopic in the field of ML, and provide the opportunity to practice skills demonstrated in the lectures in preparation for the exams. Assignments will be posted near the beginning of each such subtopic, and due within a few days of having completed each such subtopic.

Assignments will often ask for implementations of ML algorithms on real-world datasets. All such implementations must be in the Python programming language (version 3.6 or later). Assignments will involve completing cells within a Jupyter notebook skeleton. The content to be completed may be code (written in Python), plain text, or math derivations/proofs (MathJax). Please see Reference Texts for help with these languages, if they are new to you.Assignments are submitted through Crowdmark in electronic form only. For each assignment, you will submit a jupyter notebook in two forms:

- submit the .ipynb file to the D2L Dropbox, and
- submit a PDF of the jupyter notebook to Crowdmark.

It is expected that each student completes their own assignments. Discussion with other class mates is encouraged, but copying may result in a grade of zero and will be referred to the faculty as an academic integrity concern. Students will need to submit both a PDF printout of their Jupyter notebook, as well as the notebook itself so that the code can be run and verified.

As indicated in Table 2, each assignment will have a dedicated TA who can be contacted to provide clarification and help with problems. All queries about the assignment should be directed to this individual on the instructional team.

Exams

There will be both a midterm and a final exam for the course. The final exam will cover the whole course. The exams will be closed-book and no electronic devices will be allowed. Students will be allowed to bring with them one “cheat sheet”, completed with any content provided it is contained on one 8.5” x 11” sheet of paper, double-sided, into each the exams.

Project (CS 680 Only)

You are required to complete a research project, as well as provide a peer review of the proposal for one other project. Your project could involve improving state-of-the-art performance on a notable dataset, applying machine learning algorithms in an innovative way to a different field, developing a new algorithm to address a specific need in machine learning, or conducting a theoretical analysis of a machine learning algorithm (whether new or existing). Alternatively, you could do a literature survey on a trending topic in ML. You have the option to work with a partner on this project.

The project should:

- Be related to machine learning (naturally)

- Help you learn something new and engaging
- Be substantial enough to be considered for publication in a top machine learning conference

A definition of the project idea, in pitch form, is due on September 19, 2024, and will account for 5% of your final grade. This should be 3-4 sentences in length, and identify the goal of the project. These ideas are to be submitted as a discussion question on Piazza, to create the opportunity to identify students with similar interests for working in pairs.

The project proposal is due October 8. In this proposal, identify the students contributing to the project, briefly explain your project's focus, related works (no need for an exhaustive review yet), your execution plan, expected outcomes, and how you will evaluate your results. The proposal should be 2 to 4 pages long, excluding references.

The final project report is due on December 3, 2024 (the last day of classes). In this scientific report, summarize all your findings—whether empirical, algorithmic, or theoretical. Your report should include an introduction, background, main results, and a conclusion. Depending on your project, you might also include experimental results and/or a discussion section. Ensure you properly cite all prior work and results. The report should be clear, concise, and no longer than 8 pages (excluding references). Your project report will be evaluated based on clarity, significance, rigor, presentation, and completeness.

Course Policies

Office Hours

The instructor welcomes meeting students outside of course hours for scheduled meetings, either in-person or over Zoom. To meet with the instructor, please consult the calendar and generate a meeting invitation using your @waterloo.ca email in Outlook with the event name CS480-[YOUR LAST NAME]-[MEETING]. Meeting invitations must be sent at least 24 hrs prior to the proposed start time. Please indicate your preference for an in-person or virtual meeting, and if the latter, please include a Zoom link These meetings may be one-on-one or in groups of up to 3 students.Office hours with teaching assistants to discuss matters related to the assignments will be announced on the day the assignment in question is posted.

Late Work

Each student is allowed two 48-hour extensions, which can be applied to any of the four homework assignments (but only one extension per assignment). To use an extension, email both the instructor and TA at least 24 hours before the assignment's deadline, explaining your reason. You won't receive an email confirmation, but your deadline on LEARN will be adjusted at least 12 hours before the original deadline. If your deadline hasn't been updated on LEARN 12 hours before the due time despite your request, only then should you contact the course staff again. Any deviations from this process, including additional emails or Piazza posts asking about the extension status, might result in your request being ignored. Note that extensions are not available for the final project or other deadlines. Once these two extensions have been used, we will not accept late homework submissions unless there is a valid reason with documented proof (e.g., hospitalization, family emergency). Reasons like travel, being busy, or forgetting to submit are not considered valid.

Re-grading requests

You can have part of your assessment reviewed if you think we made a mistake. To do so, please email both the TA and the Instructor with "CS480/680-regrade request" in the subject title. In the body of the email, describe specifically what you think the error was with your grade. All re-grade requests must be made within 7 days of when the assessment was returned to you.

Communication Policies

Content-related matters

All content-related discussions, outside of class time or office hours, should occur on Piazza. This course has attracted students with a broad range of experience and backgrounds. As such, some topics will come more easily to some than others. To make better use of this collective expertise, Piazza will be used to facilitate discussion among students. Communications about content-related matters received through other channels (e.g. e-mail) will be directed to Piazza. We encourage all students to ask questions about the concepts covered in class on Piazza, as it is very likely that other students have the same or a similar question. As the purpose of us adopting Piazza is to foster discussion between students, the role of the instructional team will be to provide structure and guidance to the discussion. Most importantly, we want to provide space for those students who might know the answer to deepen their understanding by trying to explain it to someone else with a different perspective. The instructional team will therefore monitor Piazza, and will only respond to questions that have gone unaddressed, or to provide corrections when necessary. To engage with the instructional team directly on a content-related matter, students should make use of class time and office hours.

Logistical matters

Questions relating to logistical matters (such as late work extension requests, absences, or accommodations) should be directed via e-mail to the appropriate member(s) of the instructional team (either the TA/instructor for extensions on assignments, or the instructor for everything else). Please include CS480/680 in the subject line of the email.

Email policy

The instructional team's top priority is to provide a distillation and synthesis of the concepts that we will cover. To do this well - that is, in a way that will add something of value to students' understanding - takes time and periods of undistracted focus. Responding to emails quickly can undermine the ability to achieve this goal. Unless otherwise stated in this document, therefore, a response can be expected within ***eight (8) working hours*** of receipt of the email. The instructional team is adopting the standard 8:00am-4:00pm, Monday through Friday work day. For example, an email received at 8:00 am Monday morning can expect a response by 4:00 pm Monday afternoon.

Similarly, an email received at 4:00 pm Friday afternoon can expect a response by 4:00 pm Monday afternoon, as these are separated in time by one working minute.

Generative AI

This course includes the independent development and practice of specific skills, such as programming, and the formulation of mathematical proofs. Therefore, the use of Generative artificial intelligence (GenAI) trained using large language models (LLM) or other methods to produce text, images, music, or code, like Chat GPT, DALL-E, or GitHub CoPilot, is not permitted in this class. Unauthorized use in this course, such as running course materials through GenAI or using GenAI to complete a course assessment is considered a violation of Policy 71 (plagiarism or unauthorized aids or assistance). Work produced with the assistance of AI tools does not represent the author's original work and is therefore in violation of the fundamental values of academic integrity including honesty, trust, respect, fairness, responsibility and courage (ICAI, n.d.).

You should be prepared to show your work. To demonstrate your learning, you should keep your rough notes, including research notes, brainstorming, and drafting notes. You may be asked to submit these notes along with earlier drafts of your work, either through saved drafts or saved versions of a document. If the use of GenAI is suspected where not permitted, you may be asked to meet with your instructor or TA to provide explanations to support the submitted material as being your original work. Through this process, if you have not sufficiently supported your work, academic misconduct allegations may be brought to the Associate Dean.

In addition, you should be aware that the legal/copyright status of generative AI inputs and outputs is unclear. More information is available from the Copyright Advisory Committee:
https://uwaterloo.ca/copyright-at-waterloo/teaching/generative-artificial-intelligence

Students are encouraged to reach out to campus supports if they need help with their coursework including:

Student Success Office for help with skills like notetaking and time management
Writing and Communication Centre for assignments with writing or presentations
AccessAbility Services for documented accommodations
Library for research-based assignments

Attendance Policy

The instructional team recognizes that lectures may not necessarily provide the best learning context for all students, and expects students to use their time in whichever way maximizes the efficiency and depth of their own learning. Lecture slides will be made available on the course homepage. In-class lectures will NOT be recorded.

Mental Health Support

The Faculty of Math encourages students to seek out mental health support if needed.

On-campus Resources:
Campus Wellness https://uwaterloo.ca/campus-wellness/
Counselling Services: counselling.services@uwaterloo.ca 519-888-4567 ext. 32655
MATES: one-to-one peer support program offered by Waterloo Undergraduate Student Association (WUSA) and Counselling Services: mates@wusa.ca
Health Services: located across the creek from the Student Life Centre, 519-888-4096.

Off-campus Resources:
Good2Talk (24/7): Free confidential help line for post-secondary students. Phone: 1-866-925-5454 (Ontario and Nova Scotia only)
Here 24/7: Mental Health and Crisis Service Team. Phone: 1-844-437-3247 (Waterloo Region only)
OK2BME: set of support services for lesbian, gay, bisexual, transgender, or questioning teens. Phone: 519-884-0000 extension 213 (Waterloo Region only)

EMPOWER ME 1-833-628-5589 from Canada/US. Other countries see:
http://studentcare.ca/rte/en/IHaveAPlan_WUSA_EmpowerMe_EmpowerMe
EMPOWER ME in China: China North 108007142831

Territorial Acknowledgment

The University of Waterloo acknowledges that much of our work takes place on the traditional territory of the Neutral, Anishinaabeg and Haudenosaunee peoples. Our main campus is situated on the Haldimand Tract, the land granted to the Six Nations that includes six miles on each side of the Grand River. Our active work toward reconciliation takes place across our campuses through research, learning, teaching, and community building, and is centralized within the Office of Indigenous Relations.

University Policy

Academic Integrity

In order to maintain a culture of academic integrity, members of the University of Waterloo are expected to promote honesty, trust, fairness, respect and responsibility. See the Office of Academic Integrity webpage for more information.

Discipline

A student is expected to know what constitutes academic integrity to avoid committing an academic offence and to take responsibility for his/her actions. Check the Office of Academic Integrity for more information. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course professor, academic advisor, or the undergraduate associate dean. When misconduct has been found to have occurred, disciplinary penalties will be imposed under Policy 71 – Student Discipline. For information on categories of offenses and types of penalties, students should refer to Policy 71 - Student Discipline. For typical penalties check Guidelines for the Assessment of Penalties.

Grievance

A student who believes that a decision affecting some aspect of their university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70 - Student Petitions and Grievances, Section 4. When in doubt, please be certain to contact the department's administrative assistant who will provide further assistance.

Appeals

A decision made or penalty imposed under Policy 70 - Student Petitions and Grievances (other than a petition) or Policy 71 - Student Discipline may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 - Student Appeals.

Diversity

It is our intent that students from all diverse backgrounds and perspectives be well served by this course, and that students' learning needs be addressed both in and out of class. We recognize the immense value of the diversity in identities, perspectives, and contributions that students bring, and the benefit it has on our educational environment. Your suggestions are encouraged and appreciated. Please let us know ways to improve the effectiveness of the course for you personally or for other students or student groups. In particular:

- We will gladly honour your request to address you by an alternate/preferred name or gender pronoun. Please advise us of this preference early in the semester so we may make appropriate changes to our records.
- We will honour your religious holidays and celebrations. Please inform of us these at the start of the course.
- We will follow AccessAbility Services guidelines and protocols on how to best support students with different learning needs.

Note for Students with Disabilities

AccessAbility Services, located in Needles Hall, Room 1401, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AcessAbility Services at the beginning of each academic term.

Turnitin.com
Text matching software (Turnitin®) may be used to screen assignments in this course. Turnitin® is used to verify that all materials and sources in assignments are documented. Students' submissions are stored on a U.S. server, therefore students must be given an alternative (e.g., scaffolded assignment or annotated bibliography), if they are concerned about their privacy and/or security. Students will be given due notice, in the first week of the term and/or at the time assignment details are provided, about arrangements and alternatives for the use of Turnitin in this course.
It is the responsibility of the student to notify the instructor if they, in the first week of term or at the time assignment details are provided, wish to submit alternate assignment.

Cross-listed Course (if applicable)
Please note that a cross-listed course will count in all respective averages no matter under which rubric it has been taken. For example, a PHIL/PSCI cross-list will count in a Philosophy major average, even if the course was taken under the Political Science rubric.