CS480/680: Introduction to Machine Learning

Fall 2018

TTh 10:00 – 11:20 (MC 4021)

Web: https://cs.uwaterloo.ca/~y328yu/mycourses/480

Piazza: https://piazza.com/uwaterloo.ca/fall2018/cs480cs680/home

Syllabus

Instructor

Yao-Liang Yu
DC 3617
Assistant Professor
School of Computer Science
University of Waterloo

Web: https://cs.uwaterloo.ca/~y328yu
Email: yaoliang.yu@uwaterloo.ca
Office hours: TTh 11:30 - 12:30 or by email appointment

Course Description

This course introduces students to the design of algorithms that enable machines to "learn." In contrast to the classic paradigm where machines are programmed by specifying a set of instructions that dictate what exactly a machine should do, a new paradigm is developed whereby machines are presented with examples from which they learn what to do. This is especially useful in complex tasks such as natural language processing, information retrieval, data mining, computer vision and robotics where it is not practical for a programmer to enumerate all possible situations in order to specify suitable instructions for all situations. Instead, a machine is fed with a large collection of examples from which it automatically learns suitable rules to follow. The course will introduce the basics of machine learning and data analysis.

Course Objectives

At the end of the course, students should have the ability to:

- Recognize and formalize a task as a machine learning problem;
- Identify suitable algorithms to tackle different machine learning problems;
- Understand and implement a plethora of foundational machine learning algorithms;
- Apply machine learning algorithms to real datasets.

Course Overview

- Introduction (first lecture: Sep 6, 2018)
- Perceptron
- Linear Regression

- Multi-layer Perceptron
- Logistic Regression
- K-nearest Neighbours
- Support Vector Machines
- Kernels
- Gaussian Processes
- Mixture of Gaussians
- Deep Neural Networks
- Convolutional Neural Networks
- Recurrent Neural Networks
- Generative Adversarial Networks
- Decision Trees
- Bagging, Boosting and Random Forest
- Fairness and Privacy
- Adversarial Machine Learning

Prerequisites

Good knowledge of linear algebra (vector space, eigenvalue, matrix multiplication, etc.) and basic probability (random variable, distribution, expectation, conditional probability, Bayes rule, etc.). Exposure to numerical computing and optimization is a plus but not required. Basic programming language such as Python, Matlab, or Julia (can learn in a few hours).

Textbooks

There is no required textbook. We will pose lecture notes or slides before or after class. You are encouraged to check out the following classical books.

- *Elements of Statistical Learning (2nd edition, 2009).* Trevor Hastie, Robert Tibshirani and Jerome Friedman.
- *Deep Learning (2016).* Ian Goodfellow, Yoshua Bengio and Aaron Courville.
- *Machine Learning: A Probabilistic Perspective (2012).* Kevin Murphy.
- *Understanding Machine Learning: From Theory to Algorithms (2014).* Shai Shalev-Shwartz and Shai Ben-David.

Grading

CS480 (undergraduate students): There will be 4 homework assignments, each worth 10% of your final grade, hence in total 4 x 10% = 40%. There will be an in-class open book midterm and an open book final exam worth another 20% and 40% of your final grade, respectively. *Optionally, you can also conduct a course project to replace midterm.*

CS680 (graduate students): There will be 4 homework assignments, each worth 10% of your final grade, hence in total 4 x 10% = 40%. There will be an in-class open book midterm and an open book final exam worth another 20% and 40% of your final grade, respectively. *Optionally, you can also conduct a course project to replace final exam.*

Homework

Four homework assignments will be posted at the following dates (small variations may apply):

- assignment 1: out September 11, due September 27
- assignment 2: out September 27, due October 16
- assignment 3: out October 23, due November 8
- assignment 4: out November 31, due November 29

For students choosing to do a course project, the following due dates also apply:

- proposal: due October 23
- report: due December 3

For programming questions, choose one of the following three languages: Matlab, Python, or Julia.

Completed assignments (and project proposal and report) will be submitted through LEARN. Submit early and often!

As usual, it is OK to seek for help, but you must write your solutions independently and individually, and you should always acknowledge any help you get (book, friend, internet, etc.).

Mark appeals should be requested within two weeks of receiving the mark. The appeal could go either ways, so request only if you truly believe something is wrong.

Late Policy

We do **NOT** accept any late homework submissions, unless you have a legitimate reason with formal proof (e.g. hospitalization, family urgency, etc.). Traveling, busy with other stuff, or simply forgetting to submit, are not considered legitimate.

Project

Students can choose to conduct a research project: it could be a survey of a subfield of machine learning, or an empirical comparison of several related algorithms on an interesting dataset, or an application of machine learning algorithms to a different field, or designing a novel algorithm to address a need in machine learning, or theoretically analyzing the performance of a machine learning algorithm (new or old). Some possible projects will be suggested as we progress in the course, but you are highly encouraged to choose your own project (that interests you the most).

You project should

- relate to machine learning (obviously)
- allow you to learn something new (and hopefully significant)
- be interesting and nontrivial, preferably publishable in a top machine learning conference

The project proposal will be due on October 23. Please concisely describe what your project is about, what are the related works, what is your execution plan, what do you expect to learn/contribute, and how are you going to evaluate your results. I expect the proposal to be more or less **2 pages** (excluding references).

The project report will be due on December 3. Please summarize all your findings (empirical, algorithmic, theoretical) in a scientific report. I expect there is an introduction section, a background section, a main result section, and a conclusion section. Depending on your project, you may include an experimental section and/or discussion section. Please always give proper citations to prior work or results. Be precise and concise. I expect the report to be less than **8 pages** (everything included).

Your project report will be evaluated by its clarity, significance, rigor, presentation, and completeness.

Academic Integrity

In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. Check the university website for more information.

Grievance

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

Discipline

A student is expected to know what constitutes academic integrity to avoid committing an academic offence, and to take responsibility for hisher actions. 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 instructor, academic advisor, or the undergraduate Associate Dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline. For typical penalties check Guidelines for the Assessment of Penalties.

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).

Note for Students with Disabilities

The Office for Persons with Disabilities (OPD), located in Needles Hall, Room 1132, 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 the OPD at the beginning of each academic term.