Introduction to Machine Learning (PhD)

Spring 2019, CMU 10701

Lectures: MW, 10:30-11:50pm, Rashid Autorium: 4401 Gates and Hillman Center (GHC)
Recitations: F, 10:30-11:50pm, Rashid Autorium: 4401 Gates and Hillman Center (GHC)

Instructor: Leila Wehbe (https://www.cs.cmu.edu/~lwehbe/)

Assistant Instructor: Brynn Edmunds

Teaching Assistants:
- Ziheng Cai ()
- Byungsoo Jeon ()
- Gi Bum Kim ()
- Sarah Mallepalle ()
- Mauro Moretto ()
- Yimeng Zhang ()

Communication: Piazza (https://piazza.com/cmu/spring2019/10701) will be used for discussion about the course and assignments.

Office Hours:
- Leila: Tuesday 1.30-3.00pm, 8217 GHC
- Ziheng Cai ()
- Byungsoo Jeon ()
- Gi Bum Kim ()
- Sarah Mallepalle ()
- Mauro Moretto ()
- Yimeng Zhang ()

Course Description

Machine learning studies the question "How can we build computer programs that automatically improve their performance through experience?" This includes learning to perform many types of tasks based on many types of experience. For example, it includes robots learning to better navigate based on experience gained by roaming their environments, medical decision aids that learn to predict which therapies work best for which diseases based on data mining of historical health records, and speech recognition systems that learn to better understand your speech based on experience listening to you.

This course is designed to give PhD students a thorough grounding in the methods, mathematics and algorithms needed to do research and applications in machine learning. Students entering the class with a pre-existing working knowledge of probability, statistics and algorithms will be at an advantage, but the class has been designed so that anyone with a strong numerate background can catch up and fully participate.

If you are interested in this topic, but are not a PhD student, or are a PhD student not specializing in machine learning, you might consider the master's level course on Machine Learning, 10-601. 10-601 may be appropriate for MS and undergrad students who are interested in the theory and algorithms behind ML. You can evaluate your ability to take 10-

701 via a self-assessment exam here (https://qna-app.appspot.com/view.html?aglzfnFuYS1hcHByGQsSDFF1ZXN0aW9uTGlzdBiAgICgpO-KCgw) and see an ML course comparison here (https://docs.google.com/document/d/1Y0Jx_tcINWQrWJx31WGEQSsUs059OUMmPIVSeyxNdeM/edit).

Prerequisites

Students entering the class are expected to have a pre-existing working knowledge of probability, linear algebra, statistics and algorithms, though the class has been designed to allow students with a strong numerate background to catch up and fully participate. In addition, recitation sessions will be held to revise some basic concepts.

Resources

Supplementary readings:
[CB] *Machine Learning: A probabilistic perspective*, Kevin Murphy.
[KM] *Pattern Recognition and Machine Learning*, Christopher Bishop.
[HTF] *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Trevor Hastie, Robert Tibshirani, Jerome Friedman.
[TM] *Machine Learning*, Tom Mitchell.

Schedule

Tentative schedule, might change according to class progress and interest. Every Friday classes is intended to be a recitation to review material or answer homework questions, however this might change if we need a makeup lecture.

Date	Note	Topic	Resources
		Basics	
01/14		Lecture 1: Introduction - What is Machine Learning	
01/16	HW1 Out	Lecture 2: Building blocks - MLE, Bayesian Estimation, MAP, Decision Theory, Model-free, Risk Minimization	
01/18	Special Office Hours		
01/21	MLK day, no class		
01/23		Lecture 3: Classification, kNN	
01/25		Recitation	
		Parametric Estimation and Prediction	
01/28		Lecture 4: Linear Regression, Regularization	
01/30	HW1 due, HW2 Out	Lecture 5: Logistic Regression ()	
02/01		Recitation ()	
02/04		Lecture 6: Naive Bayes ()	
02/06		Lecture 7: Disciminative vs generative ()	
02/08		Recitation ()	
02/11		Lecture 8: Decision Trees ()	
02/13	HW2 due, HW3 Out	Lecture 9: Neural Networks (perceptron, neural nets) ()	
02/15		Recitation ()	
02/18	Project Topic Selection	Lecture 10: Neural Networks (deep nets, backprop) ()	
02/20		Lecture 11: SVMs ()	
02/22		Recitation ()	

02/25	Course Drop Deadline	Lecture 12: SVMs ()
02/27	HW3 due	Lecture 13: Boosting ()
03/01		Recitation ()

Learning Theory

03/04		Lecture 14: Generalization, Model selection ()
03/06	Midway Report Due	Lecture 15: Learning Theory ()
03/08	Mid-Semester Break, no class	
03/11	Spring break, no class	
03/13	Spring break, no class	
03/15	Spring break, no class	

Unsupervised Learning

| 03/18 | | Lecture 16: Clustering () |
| 03/20 | | Lecture 17: Clustering () |

| 03/21 | | |

Midterm Exam (Thursday 3/21 6:30pm)

03/22		Recitation ()
03/25	HW4 Out	Lecture 18: Representation Learning ()
03/27		Lecture 19: Representation Learning ()
03/29		Recitation ()

Graphical Models and Structured Prediction

04/01		Lecture 20: Graphical Models ()
04/03		Lecture 21: Graphical Models ()
04/05		Recitation ()
04/08	HW4 Due	Lecture 22: Graphical Models ()
04/10		Lecture 23: Sequence Models ()
04/12	Spring Carnival, No Class	
04/15		Lecture 24: Sequence Models ()

Special Topics

04/17		Lecture 25: Semi-supervised learning ()
04/19		Recitation ()
04/22		Lecture 26: Online learning ()
04/24		Lecture 27: Reinforcement learning ()
04/26		Recitation ()
04/29		Lecture 26: Deep Reinforcement learning ()
05/01	Final Reports Due	Lecture 27: Review, discussion ()
05/03	Last day of class	()

Grading

The final grade will be determined as follows:

| Exam 1 | 15% |
| Exam 2 | 15% |

Homework	40%
Project	25%
Participation	5%

Course Policies

(The following policies are adapted from and Ziv Bar-Joseph and Pradeep Ravikumar 10-701 Fall 2018 (http://www.cs.cmu.edu/~pradeepr/701/) and Roni Rosenfeld's 10-601 Spring 2016 (http://www.cs.cmu.edu/~roni/10601/) Course Policies.)

Homework

There will be four homework assignments, each worth 10% of the final grade. Answers will be submitted on Gradescope (through canvas) and code portions through Autolab.

Projects

You may work in teams of 3-5 people. There will be a limited number of project to choose from. You will not be able to chose other projects. There will be two deliverables (project proposal and final report). Each team member's contribution should be highlighted. You should use the project as an opportunity to "learn by doing".

Extensions

You will have 8 late days (for the entire semesters) that you can use for homework submission. You can chose to divide the days up on the homework assignments the way you want, with the constraint that you cannot use more than 4 days for any given homework. After you exceed your allowed days, you will lose all of the points for a late homework. You should use these dates if you need extensions for various deadlines (conferences, interviews etc.). If for some reason you need extension for some other reason, you should talk to us at least five days in advance. In the case of a true emergency that cannot be predicted (sickness, family problems etc.) we can give a reasonable extension. You cannot use late days for the projects.

Academic Integrity

Collaboration among Students

Collaboration among students is allowed but is intended to help you learn better. You can work on solving assignments together, but you should always write up your solutions separately. You should always implement code alone as well. Whenever collaboration happens, it should be reported by all parties involved in the relevant homework problem.

Online Resources

Some of the homework question you receive might have solutions online. Looking up the answers to these homework questions is not allowed. Similarly, looking up code for a problem is not allowed. Sometimes, you might need help with a small portion of the code (for example array indexing), such basic things that are not in relation to understanding the material are allowed. Do use office hours if you have questions about homework problems.

Disclosing

Whenever you collaborate with someone or you look at online material, you should disclose it in your homework. When in doubt, always disclose. Any breach of academic integrity policies will be reported to the university authorities (your Department Head, Associate Dean, Dean of Student Affairs, etc.) as an official Academic Integrity Violation and will carry severe penalties. It is truly both an unethical and a bad strategy.

Accommodations for Students with Disabilities

If you have a disability and are registered with the Office of Disability Resources, I encourage you to use their online system to notify me of your accommodations and discuss your needs with me as early in the semester as possible. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, I encourage you to contact them at access@andrew.cmu.edu..

Take care of yourself

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/ (http://www.cmu.edu/counseling/). Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:
- CaPS: 412-268-2922
- Re:solve Crisis Network: 888-796-8226

If the situation is life threatening, call the police
- On campus: CMU Police: 412-268-2323
- Off campus: 911