Syllabus
CS 6923 Machine Learning
Fall 2018

Instructor: Prof. Lisa Hellerstein
Office: 2 Metrotech, Rm. 10.092
Phone: 646-997-3689
Email: lisa.hellerstein@nyu.edu
Office Hours: Wed. 2:30-4:30 p.m.

TEXTBOOKS

Required:

The main textbook for this course is *Introduction to Machine Learning*, by Ethem Alpaydin, Third Edition. Published by MIT Press. (available online through the NYU library). It covers many of the topics in the course clearly, but is terse.

Optional/Recommended:

For a more intuitive and accessible approach to some of the topics in the course, see

1. Daume: *A Course in Machine Learning* by Hal Daume (unfinished book draft), `http://ciml.info/`

2. WHT: *An Introduction to Statistical Learning with Applications* in R by James, Witten, Hastie, and Tibshirani. `http://www-bcf.usc.edu/~gareth/ISL/`

3. Mitchell: Machine Learning, by Tom Mitchell. Published by McGraw Hill.

4. WFH: *Data Mining: Practical Machine Learning Tools and Techniques*, by Witten, Frank, and Hall. Third Edition. Published by Morgan Kauffmann.

The following books are more comprehensive than the Alpaydin text and assume that the reader has a stronger math/statistics background:

1. HTF: Hastie, Tibshirani, Friedman, Elements of Statistical Learning, Second Edition Published by Springer. (available online through the NYU library)

2. Bishop: Pattern Recognition and Machine Learning, by Christopher Bishop. Published by Springer.

Homework:

Homeworks will include written exercises as well as hands-on work with datasets and tools. All programming must be done in MATLAB (or the free package Octave) or Python. We will provide Matlab information for students who want to learn it.

Exams:

The course will have a midterm and a final exam. Academic Honesty: Students may discuss homework with other students, but must write up their own solutions in their own words, and do their own coding. DO NOT TAKE THIS COURSE IF YOU DO NOT THINK YOU CAN DO THE WORK YOURSELF, HONESTLY. See below for NYU School of Engineering Conduct Policy.

Prerequisites:

CS5403, CS6003, and CS6033 or equivalent. Exceptions are by permission of the instructor.

Note:

Machine Learning is more mathematical than most other graduate CS courses. Students often have difficulty with this course (and risk getting a grade of C or lower) if they know how to program but have not taken much math, or have not done well in their math courses. You should have taken an undergraduate course that covered probability and statistics. You should know about the probability density function (pdf), cumulative density function (cdf), continuous probability distributions, conditional probability, and expected values. You should know about partial derivatives and gradients (or be prepared to learn about them on your own). You should also know the basics of linear algebra.

Grading:

The midterm will count 25% of your grade, the final 40%, and the homeworks/project 35%.

Tentative Schedule:

We may not get through all the material in this schedule.

Lecture	Date	Topics	Reading
1	1/24	Introduction, Bayesian Decision Theory	Alpaydin:3 HTF: 1
2	1/31	Naive Bayes, k-NN, ML Experiments	Alpaydin: 3,19 Mitchell online: see below*
3	2/7	MLE, Bias/Variance	Alpaydin:Chap. 4
4	2/14	Linear and Logistic Regression	Alpaydin: Chap 4 Mitchell online: see below* Daume: Chap. 2
5	2/21	Decision Trees and Random Forests	Alpaydin: Chap. 9
6	2/28	Data Preprocessing, Dimensionality Reduction	Alpaydin: Chap. 6
7	3/7	Computational Learning Theory	
8	3/21	Midterm	
9	3/28	Perceptrons, Neural Nets, Deep Learning	Alpaydin:Chap. 11
10	4/4	Kernel Machines, SVMs	Alpaydin:Chap. 13
11	4/11	Clustering	Alpaydin:Chap. 7
12	4/18	Ensemble Methods	
13	4/25	Graphical Models	Alpaydin:Chap. 16
14	5/2	Reinforcement Learning	Alpaydin:Chap. 18
15	5/9	FINAL EXAM	

*Mitchell on-line draft chapter on Naive Bayes and Logistic Regression: `https://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf`

Moses Center Statement of Disability:

If you are student with a disability who is requesting accommodations, please contact New York University's Moses Center for Students with Disabilities (CSD) at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations.

Information about CSD can be found at `www.nyu.edu/csd`. It is located at 726 Broadway on the 2nd floor.

NYU School of Engineering Policies and Procedures on Academic Misconduct

Introduction:

The School of Engineering encourages academic excellence in an environment that promotes honesty, integrity, and fairness, and students at the School of Engineering are expected to exhibit those qualities in their academic work. It is through the process of submitting their own work and receiving honest feedback on that work that students

may progress academically. Any act of academic dishonesty is seen as an attack upon the School and will not be tolerated. Furthermore, those who breach the School's rules on academic integrity will be sanctioned under this Policy. Students are responsible for familiarizing themselves with the School's Policy on Academic Misconduct.

Definition:

Academic dishonesty may include misrepresentation, deception, dishonesty, or any act of falsication committed by a student to inuence a grade or other academic evaluation. Academic dishonesty also includes intentionally damaging the academic work of others or assisting other tudents in acts of dishonesty. Common examples of academically dishonest behavior include, but are not limited to, the following:

1. Cheating: intentionally using or attempting to use unauthorized notes, books, electronic media, or electronic communications in an exam; talking with fellow students or looking at another person's work during an exam; submitting work prepared in advance for an in-class examination; having someone take an exam for you or taking an exam for someone else; violating other rules governing the administration of examinations.

2. Fabrication: including but not limited to, falsifying experimental data and/or citations.

3. Plagiarism: intentionally or knowingly representing the words or ideas of another as one's own in any academic exercise; failure to attribute direct quotations, paraphrases, or borrowed facts or information.

4. Unauthorized collaboration: working together on work that was meant to be done individually.

5. Duplicating work: presenting for grading the same work for more than one project or in more than one class, unless express and prior permission has been received from the course instructor(s) or research adviser involved.

6. Forgery: altering any academic document, including, but not limited to, academic records, admissions materials, or medical excuses.

See `http://engineering.nyu.edu/files/SACCofC9-14-15.pdf` for the Code of Conduct, which can be found here: `http://engineering.nyu.edu/life/student-affairs/community-standards-procedures`