CS281: Advanced Machine Learning (Fall 2017)

Instructor:	Prof. Alexander (Sasha) Rush
Teaching Fellows:	Rachit Singh, Mark Goldstein, Yuntian Deng, and Carl Denton
Lectures:	Monday and Wednesday, 1-2:30pm
Location:	Maxwell-Dworkin G-115
Section:	Fri, 4-5: MD 119
URL:	`http://cs281.fas.harvard.edu/`

This Syllabus

This document reviews the course description, prerequisites, major components of the course, and course policies. All information about the lectures, sections, office hours, assignments, textbook, readings, course grading, and Piazza discussion board may also be found on the course website.

Course Description

Advanced statistical machine learning and probabilistic data analysis. Topics include: Variational inference, graphical models, deep learning, text modeling, unsupervised learning, dimensionality reduction and visualization. Requires a major final project.

This course is about learning to extract statistical structure from data, for making decisions and predictions, as well as for visualization. The course will cover many of the most important mathematical and computational tools for probabilistic modeling, as well as examine specific models from the literature and examine how they can be used for particular types of data.

There will be a heavy emphasis on implementation. All staff-provided code will use Python. The course will make heavy use of NumPy and the PyTorch library. You are free to use another language of your choice, but the staff may not be able to offer much assistance. All of the homework assignments will require some amount of coding, and the final project will certainly require the running of computational experiments.

Prerequisites

You should feel comfortable with the basics of probability: joint densities, conditional distributions, etc. An undergraduate level is fine — there will not be much measure theory and such background is not required. There will be linear algebra, e.g., solving linear systems and thinking about eigenvalues/eigenvectors. We will regularly use multivariable calculus. Additionally, prior exposure to machine learning on the level of CS181 is highly recommended.

Readings and Lecture Preparation

Each week will have some required reading, typically from the Murphy text and some occasional papers. There will also be optional reading and videos to watch. Machine learning is unique in that there are many videos that are freely available on the web of world-leading researchers discussing their work. It is highly recommended to watch these videos. Everyone learns in a different way and so it may be helpful to hear the same material presented from a slightly different point of view. Moreover, the availability of these videos means that class time can be used for dynamic interaction rather than just lecturing.

There will be a strongly-recommended weekly section led by the teaching staff. The material covered in section will provide examples and applications of the material presented in lecture and will be pertinent to the assignments and to the final projects; it is highly recommended that you attend section.

Assignments

There will be five assignments due over the course of the term. Please see the course website for the homework schedule.

The assignments should be completed using LaTeX; the template file will be available with the problem set. All assignments should be submitted electronically by 5:00pm on the due date via Canvas. It is your responsibility to ensure that the PDF files are readable.

Collaboration Policy You are encouraged to discuss high-level solutions and problem-solving approaches with other students in the class, but ultimately you must write your own code and produce your own results. If you have collaborated with other students in the planning and design of solutions to the assignments, provide their names on your writeup in the space provided.

Self-Grading Policy Once the submission deadline has passed, students will receive the answer key for the assignment. Students will then be responsible for using this answer key to submit their own scores; this process is there to ensure that students check their understanding and to provide very rapid feedback. The staff will spot-check submitted scores for accuracy. *It is a honor code violation to look at the answer key if you haven't yet turned in your assignment (e.g. using late days). It is also an honor code violation to look at an answer key from a previous iteration of the class.*

Regrading Policy If you think there is a mistake with the grading of your midterm or one of your assignments, you may submit it to the course staff for a regrade. Note that this regrade is a new draw from the grading distribution and may result in your grade moving up or down. There is a time limit on these requests: you may only ask for a regrade within two weeks of the section at which the assignment/midterm was returned.

Late Policy You have a total of 5 late days that can be used for assignments. Weekends count as late days. Please use your late days carefully. Late days cannot be applied to the final project write-up.

Midterm Exam

There will be an in-class midterm exam, scheduled for October 23, 2017. This exam will cover material from the beginning of the class, through the October 16th lecture on Mixture Models.

Final Project

In the second half of the course, you will complete a project. The ideal outcome of this project would be a paper that could be submitted to a top-tier machine learning conference such as NIPS, ICML, UAI, AISTATS, or KDD. There are different ways to approach this project, which are discussed in a more comprehensive document that is available on the course website. There are four separate components of the project:

Proposal, Due Oct 13th This is a two-page document that describes the problem you intend to solve, your approach to solving it and the experiments that you intend to run to evaluate your solution.

Abstract and Status Report, Due Nov 3rd This is a three to four page document that contains a draft of your final abstract, as well as a brief status report on the progress of your project.

Poster Session, Dec 1st You will make a conference-style poster about your project, with a class poster session during reading week. SEAS will pay for the cost of producing the poster. You will also submit a PDF file of your poster.

Final Report, Due Dec 8th You will write a report of up to ten pages, in the style of a mainstream CS conference paper. See the papers linked on the course website for examples.

You will upload these materials via Canvas. Please see the document referenced above (linked in the course website) for a more thorough description of the final project and policies related to collaboration, etc.

Grading

Please see the Grading section of the course website for more details.

Textbooks and References

The following book is required for the course:

Machine Learning: A Probabilistic Perspective Kevin P. Murphy, MIT Press, 2012. This book covers a wide set of important topics. As the book is fresh and comprehensive, there are still quite a few errors. We will try to maintain lists of errata as they are discovered. Please see the course website for the reading schedule.

The following book is strongly recommended, but not required:

Pattern Recognition and Machine Learning Christopher M. Bishop, Springer, 2006. An excellent and affordable book on machine learning, with a Bayesian focus. It covers fewer topics than the Murphy book, but goes into greater depth on many of them and you may find that you prefer Bishop's exposition.

These are other (free online!) books on machine learning and related topics that you may find helpful, but that are completely optional.:

Information Theory, Inference, and Learning Algorithms David J.C. MacKay, Cambridge University Press, 2003. Freely available online at `http://www.inference.phy.cam.ac.uk/mackay/itila/`. A very well-written book with excellent explanations of many machine learning topics.

Bayesian Reasoning and Machine Learning David Barber, Cambridge University Press, 2012. Freely available online at `http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.Online`.

The Elements of Statistical Learning Trevor Hastie, Robert Tibshirani, and Jerome Friedman, Springer, 2009. Freely available online at `http://www-stat.stanford.edu/~tibs/ElemStatLearn/`.

These are books on some specialized topics that you may find useful:

Gaussian Processes for Machine Learning Carl Edward Rasmussen and Christopher K.I. Williams, MIT Press, 2006. Freely available online at `http://www.gaussianprocess.org/gpml/`.

Non-Uniform Random Variate Generation Luc Devroye, Springer-Verlag, 1986. Freely available online at `http://luc.devroye.org/rnbookindex.html`.

Probabilistic Graphical Models: Principles and Techniques Daphne Koller and Nir Friedman, MIT Press, 2009.

Numerical Optimization Jorge Nocedal and Stephen J. Wright, Springer, 2006.

Bayesian Data Analysis Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. CRC, 2013.

Elements of Information Theory Thomas M. Cover and Joy A. Thomas, Wiley, 1991.

Monte Carlo Statistical Methods Christian P. Robert and George Casella, Springer, 2005.

Piazza

Most questions about the course, lecture or section material, or the assignments should be addressed via

Piazza at `https://piazza.com/harvard/fall2017/cs281`. The course instructors will check this discussion board and make an effort to post responses within a day. Students taking the class are also encouraged to post responses. Code examples can be posted, but don't post anything you wouldn't be expected to share with other students in the class as per the collaboration policy. Long, detailed questions are probably best answered during office hours. Questions that are not appropriate for the discussion board may be sent to instructors privately through Piazza. Use your judgement. **Note that setting a public post as anonymous will hide your name from classmates. Instructors will always see your name on a public, anonymous public, or private post.**