CS281: Advanced Machine Learning (Fall 2015)

Instructor:	Prof. Finale Doshi-Velez
Teaching Fellows:	David Duvenaud, José Miguel Hernandez-Lobato, Yakir Reshef, Keyon Vafa
Lectures:	Monday and Wednesday, 9-10:30am
Location:	60 Oxford Street 330
Section:	TBD
Office Hours:	Monday 11:15-12:30pm in MD 219 (Doshi-Velez)
	Monday 8-9pm in Leverett Dining Hall (Vafa)
	Wednesday 10:30-11:30am in MD Floor 2 Lobby (Reshef)
	Friday 2:30-3:30pm in MD Floor 2 Lobby (Duvenaud)
	TBD (Hernandez-Lobato)
URL:	`https://canvas.harvard.edu/courses/6058`

Course Description

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. You may use any language that you wish, but the staff will provide support only for Python. Four of the five assignments will involve some amount of coding, and the final project will almost certainly require the running of computer 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 some basic linear algebra, e.g., solving linear systems and thinking about eigenvalues/eigenvectors. We will regularly use multivariable calculus.

Readings and Lecture Preparation

Each week will have some required reading, typically from books and 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 be extremely pertinent to the assignments and to the final projects; it is highly recommended that you attend section.

In-Class Quizzes

To help encourage you to do the reading, there will be in-class quizzes at the beginning of most lectures. These will be a very simple, multiple choice quizzes that are not intended to be difficult if you have read the required material. These quizzes will comprise 10% of your final grade.

Assignments

There will be five assignments due over the course of the term. Each of these will be worth 7% of your final grade. These will be assigned and due according to the following schedule, in which you have two weeks for each one:

	Out	*Due*
Assignment 0	2 September 2015	11 September 2015
Assignment 1	9 September 2015	23 September 2015
Assignment 2	23 September 2015	7 October 2015
Assignment 3	7 October 2015	28 October 2015
Assignment 4	28 October 2015	11 November 2015

The assignments should be completed using LaTeX; the template file will be available with the problem set. All assignments should be submitted electronically by 23:59 on the due date via Canvas. It is your responsibility to ensure that the PDF files are readable. We will make every effort to grade the assignments promptly and return them to you in the next section.

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.

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 four late days that can be used for assignments 0-4 and your project proposal and update. 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 Wednesday 21 October 2015. This exam comprises 15% of your grade and will cover material from the beginning of the class, through the Friday 14 October lecture on Variational Inference.

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 from the course website under the Files tab. There are four separate components of the project:

Proposal, Due 14 October 2015 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. The proposal represents 10% of your overall course grade.

Abstract and Status Report, Due 6 November 2015 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. This represents 5% of your overall course grade.

Poster Session, 4 December 2015 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. This will represent 5% of your course grade.

Final Report, Due 11 December 2015 You will write a report of up to ten pages, in the style of a mainstream CS conference paper. This will represent 20% of your overall grade.

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

Grading

Each Assignment	7%
In-Class Quizzes	10%
Midterm	15%
Final Project Proposal	10%
Final Project Abstract and Status Report	5%
Final Project Poster	5%
Final Project Report	20%

Textbooks and References

The following book is required for the course:

Machine Learning: A Probabilistic Perspective Kevin P. Murphy, MIT Press, 2012. This is a very new book that 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.

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.

Discussion Board

Most questions about the course, lecture or section material, or the assignments should be addressed via Piazza at `https://piazza.com/harvard/fall2015/compsci281`. 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.

Anticipated Schedule

2 Sep 2015 Lecture 1: Introduction to Inference and Learning

4 Sep 2015 Section: Math Review

7 Sep 2015 Lecture 2: Simple Discrete Models

9 Sep 2015 Lecture 3: Simple Gaussian Models; Assignment 0 Due

11 Sep 2015 Section: Code Review

14 Sep 2015 Lecture 4: Bayesian Statistics

16 Sep 2015 Lecture 5: Linear Regression

18 Sep 2015 Section: Practical Optimization

21 Sep 2015 Lecture 6: Linear Classifiers

23 Sep 2015 Lecture 7: Generalized Linear Models; Assignment 1 Due

25 Sep 2015 Section: Undirected Graphical Models and Factor Graphs

28 Sep 2015 Lecture 8: Directed Graphical Models

30 Sep 2015 Lecture 9: Mixture Models

2 Oct 2015 Section: Sparse Linear Models

5 Oct 2015 Lecture 10: Factor Analysis and PCA

7 Oct 2015 Lecture 11: Exact Inference; Assignment 2 Due

9 Oct 2015 Section: The Junction Tree Algorithm

14 Oct 2015 Lecture 12: Variational Inference; Project Proposal Due

16 Oct 2015 Section: Loopy Belief Propagation and Expectation Propagation

19 Oct 2015 Lecture 13: Monte Carlo Basics

21 Oct 2015 In-Class Midterm

23 Oct 2015 Section: Particle Filtering

26 Oct 2015 Lecture 14: Markov Chain Monte Carlo

28 Oct 2015 Lecture 15: Advanced MCMC; Assignment 3 Due

30 Oct 2015 Section: MCMC Practicalities

2 Nov 2015 Lecture 16: Spectral Methods

4 Nov 2015 Lecture 17: Latent Dirichlet Allocation

6 Nov 2015 Section: Practical Spectral Methods; Project Update Due

9 Nov 2015 Lecture 18: Time Series Models I

11 Nov 2015 Lecture 19: Time Series Models II; Assignment 4 Due

13 Nov 2015 Section: Practical Time Series Models

16 Nov 2015 Lecture 20: Kernels and Gaussian Processes

18 Nov 2015 Lecture 21: Dirichlet Processes I

20 Nov 2015 Section: Practical Gaussian Processes;

23 Nov 2015 Lecture 22: Dirichlet Processes II

30 Nov 2015 Lecture 23: Neural Networks

2 Dec 2015 Lecture 24: Advanced Neural Networks

4 Dec 2015 Final Project Poster Session (Tentative)

11 Dec 2015 Final Project Reports Due