Syllabus
Probabilistic Models of
Human and Machine Intelligence
CSCI 5822/ Spring 2018
Tu, Th 11:00-12:15
ECCR 105

Instructor
Professor Michael Mozer

Course Objectives

For humans and machines, intelligence requires making sense of the world---inferring simple explanations for the mishmosh of information coming in through our senses, discovering regularities and patterns, and being able to predict future states. In artificial intelligence and cognitive science, the formal language of probabilistic reasoning and statistical inference have proven useful to model intelligence. From a probabilistic perspective, knowledge is represented as degrees of belief, observations provide evidence for updating one's beliefs, and learning allows the mind to tune itself to statistics of the environment in which it operates.

One virtue of probabilistic models is that they straddle the gap between cognitive science, artificial intelligence, and machine learning. The same methodology is useful for both understanding the brain and building intelligent computer systems. Indeed, for much of the research we'll discuss, the models contribute both to machine learning and to cognitive science. Whether your primary interest is in engineering applications of machine learning or in cognitive modeling, you'll see that there's a lot of interplay between the two fields.

The course participants are likely to be a diverse group of students, some with primarily an engineering/CS focus and others primarily interested in cognitive modeling (building computer simulation and mathematical models to explain human perception, thought, and learning).

Prerequisites

The course is open to any students who have some background in cognitive science or artificial intelligence and who have taken an introductory probability/statistics course or the graduate machine learning course (CSCI 5622). If your background in probability/statistics is weak, you'll have to do some catching up with the text.

Course Readings

We will be using the text Bayesian Reasoning And Machine Learning by David Barber (Cambridge University Press, 2012). The author has made available an electronic version of the text. Note that the electronic version is a 2015 revision. Because the electronic version is more recent, all reading assignments will refer to section numbers in the electronic version.

For additional references, wikipedia is often a useful resource. The pages on various probability distributions are great references. If you want additional reading, I recommend the following texts:

- Pattern Recognition and Machine Learning by Chris Bishop
- Bayesian Data Analysis by Gelman, Carlin, Stern, & Rubin
- Machine Learning: A Probabilistic Perspective by Kevin Murphy [be sure to get the fourth printing; there were many typos in earlier versions]
- Bayesian cognitive modeling: A practical course by Michael Lee and Erik-Jan Wagenmakers [electronic version online]
- We will also be reading research articles from the literature, which can be downloaded from the links on the class-by-class syllabus below.

Course Discussions

We will use Piazza for class discussion. Rather than emailing me, I encourage you to post your questions on Piazza. Feel free to post anonymously. I strive to respond quickly. If I do not, please email me personally. To sign up, go here. The class home page is here.

Course Requirements

Readings

In the style of graduate seminars, your will be responsible to read chapters from the text and research articles before class and be prepared to come into class to discuss the material (asking clarification questions, working through the math, relating papers to each other, critiquing the papers, presenting original ideas related to the paper).

Homework Assignments

We can all delude ourselves into believing we understand some math or algorithm by reading, but implementing and experimenting with the algorithm is both fun and valuable for obtaining a true understanding. Students will implement small-scale versions of as many of the models we discuss as possible. I will give about 10 homework assignments that involve implementation over the semester, details to be determined. Most students in the class will prefer to use python, and the tools we'll use are python based. If you have a strong preference, matlab is another option. For one or two assignments, I'll ask you to write a one-page commentary on a research article.

Semester Grades

Semester grades will be based 5% on class attendance and participation and 95% on the homework assignments. I will weight the assignments in proportion to their difficulty, in the range of 5% to 15% of the course grade. Students with backgrounds in the area and specific expertise may wish to do in-class presentations for extra credit.
Procedures for Homework Assignments

General Policy

You may work either individually or in a group of two. If you work with someone else, I expect a higher standard of work. I'm not proud to tell you this, but from 30 years of grading, I have to warn you that professors and TAs have a negative predisposition toward hand printed work. It is much easier to digest responses that are typed, spell corrected, and have made an effort to communicate clearly. We will be grading not only on the results you obtain but on the clarity of your write up.
Because of the large class size, no late assignments will be accepted without a medical excuse or personal emergency. If you have a conflicting due date in another class, give us a heads-up early and we'll see about shifting the due date.

Instructor and TA are eager to help folks who are stuck or require clarification. For any clarification of the assignment, what we're expecting, and how to implement, we would appreciate it if you post your question on piazza. In fact, post on piazza unless your question is personal or you believe it is specific to you. If you have the question, it's likely others will have the same question. And if we give you a clue, then we'll give the same clue to everyone else.
See additional information at the end of the syllabus on academic honesty.

Submission of Work

We ask you to submit a hardcopy of your write up (but not code) in class on the due date.
We also ask that you upload your write up and any code as a .zip file on moodle (instructions below).
Be sure to write your full name on the hardcopy and in the code.
If you are working in a group, hand in only one hard copy and put both of your names on the write up and code.
We ordinarily will not look at your code, unless there appears to be a bug or other problem.

To submit on moodle:
(1) go to moodle.cs.colorado.edu and enter identikey and password
(2) select CSCI 5822 (key CSCI5822-S18)
(3) search for the assignment number and open the link
(4) Click on the "add submission" button
(5) Upload the .zip file containing write up and code.

Class-By-Class Plan and Course Readings

I've done my best to plan the whole semester, but we will have to revise as we go along. Take any part of this schedule as tentative if it is more than 2 weeks out.

Date	Activity	Required Reading (Section numbers refer to 2015 edition of Barber)	Optional Reading	Lecture Notes	Assignments
Jan 16	introductory meeting	Appendix A.1-A.4, 13.1-13.4	Chater, Tenenbaum, & Yuille (2006)	lecture	Assignment 0
Jan 18	basic probability, Bayes rule	1.1-1.5, 10.1	Griffiths & Yuille (2006)	lecture	
Jan 23	continuous distributions	8.1-8.3		lecture	Assignment 0 due
Jan 25	concept learning, Bayesian Occam's razor	12.1-12.3 (omit 12.2.2, which requires some probability we haven't yet talked about)	Tenenbaum (1999) Jefferys & Berger (1991)	lecture	Assignment 1
Jan 30	Gaussians	8.4-8.5		Lecture	
Feb 1	motion illusions as optimal percepts	Weiss, Simoncelli, Adelson (2002)	motion demo 1 motion demo 2	lecture	Assignment 2
Feb 6	<catch up day>			lecture	
Feb 8	Bayesian statistics (conjugate priors, hierarchical Bayes)	9.1	useful reference: Murphy (2007)	lecture	
Feb 13	Bayes nets: Representation	2.1-2.3, 3.1-3.5	Cowell (1999) Jordan & Weiss (2002) 4.1-4.6	lecture	Assignment 3
Feb 15 & Feb 20	Bayes nets: Exact Inference	5.1-5.5	Huang & Darwiche (1994)	lecture	
Feb 22 & Feb 27	Bayes nets: Approximate inference	27.1-27.6	Andrieu et al. (2003)	lecture	Assignment 4
Mar 1					Assignment 5
Mar 6	Learning I: Parameter learning GUEST: Antonio Blanca	8.6, 9.2-9.4	Heckerman (1995) 9.5	Lecture	
Mar 8, 13	Learning II: Missing data, latent variables, EM, variational methods	11.1-5, 20.1-3 28.1-28.5 28.6-28.9		Lecture	
Mar 15	Learning III: learning model structure GUEST: Andrew Lan		Lan (2018)	lecture	Assignment 6
Mar 20	text mining latent Dirichlet allocation	20.6	Griffiths, Steyvers & Tenenbaum (2007) Blei, Ng, & Jordan (2003) video tutorial on Dirichlet Processes by Teh or Teh introductory paper	Lecture	
Mar 22	text mining topic model extensions		McCallum, Corrado-Emmanuel, & Wang (2005) Bamman, Underwood, & Smith (2014)	lecture	Assignment 7

Apr 3, 5	nonparametric Bayes hierarchical models		Orbanz & Teh (2010) Teh (2006)	lecture1 lecture2	Assignment 8
Apr 10	modeling and optimization Gaussian processes		Shahriari, Swersky, Wang, Adams, and de Freitas	lecture	Assignment 7 due
Apr 12,17	modeling and optimization Multiarm bandits and Bayesian optimization			lecture	Assignment 8 due
Apr 19	sequential models hidden Markov models conditional random fields	23.1-23.5	Gharamani (2001) Sutton & McCallum Mozer et al. (2010) Lafferty, McCallum, Pereira (2001)	lecture 1 lecture 2	Assignment 9
Apr 24	sequential models Kalman filters	24.1-24.4	Koerding, Tenenbaum, & Shadmehr (2007) 24.5	lecture	
Apr 26	sequential models exact and approximate inference (particle filters, changepoint detection)	27.6 Adams & MacKay (2008) Yu & Cohen (2009)	Wilder, Jones, & Mozer (2010)	lecture1 lecture2	
May 1,3	probabilistic models and deep learning			lecture1 lecture2	Assignment 9 due (May 3)
Wed May 9, 16:30-19:00	Reserve for possible final project presentations				

Queue
Peter Welinder, Steve Branson, Serge Belongie, Pietro Perona
The Multidimensional Wisdom of Crowds

The Wisdom of Crowds in the Recollection of Order Information (2009)
Mark Steyvers, Michael Lee, Brent Miller, Pernille Hemmer

Interesting Links

Tutorials

- Owen Lewis's review of probabilistic models
- Josh Tenenbaum's Bayesian models tutorial at NIPS 2006
- Carl Rasmussen's Gaussian Processes tutorial at NIPS 2006
- Jordan tutorial on hierarchical Dirichlet processes
- Andrew Moore's tutorials
- Inference in belief networks: A procedural guide (Huang & Darwiche, 1994)
- Bayesian inference with tears (Kevin Knight) -- particularly useful for those interested in NLP
- video lecture from summer school
- Rob Lindsey's notes on basic probability and statistics

Modeling tools

- UCI Topic modeling toolbox (requires 32-bit matlab)
- Mallet (machine learning for language, Java based implementation of topic modeling)
- Mahout (Java API that does topic modeling)
- C implementatoni of topic models
- windows executable of C implementation (runs from the command line)
- Stanford Topic Modeling Toolkit
- UCLA's samiam
- Bayesian nonparametric machine learning for python
- Edward - Bayesian statistics, machine learning, probabilistic programming, deep learning (manual)
- Pyro - deep universal probabilistic programming

- PNL - probabilistic networks library
- pyMC3 - probabilistic programming with Theano
- Murphy's probabilistic modeling toolbox
- OpenBUGS - Bayesian inference using Gibbs sampling
- Dimple
- Bayesian reasoning and machine learning software in matlab (associated with David Barber's book)
- Chris DeHoust comments on software
- Augur (may not yet be available)