Syllabus
Probabilistic Models of Human and Machine Intelligence
CSCI 7222
Fall 2015

Tu, Th 11:00-12:15
Muenzinger D430

Instructor

Professor Michael Mozer
Department of Computer Science
Engineering Center Office Tower 741
(303) 492-4103
Office Hours: Th 13:00-14:30

Course Objectives

A dominant paradigm in artificial intelligence and cognitive science views the mind as a computer extraordinarily well tuned to statistics of the environment in which it operates. From this perspective, learning--by both humans and machines--involves collecting observations and updating statistics. The goal of the course is to understand the latest advances in theory in artificial intelligence and cognitive science that take a statistical and probabilistic perspective on learning and intelligence.

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. 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 a text by David Barber (Bayesian Reasoning And Machine Learning, 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:

- Chris Bishop's Pattern Recognition and Machine Learning
- Bayesian Data Analysis by Gelman, Carlin, Stern, & Rubin
- Machine Learning: A Probabilistic Perspective by Kevin Murphy (which is being used in CSCI 5622 this semester)

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. 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. My preference is for you to work in matlab, both because you can leverage software available with the Barber text, and because matlab has become the de facto work horse in machine learning. 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 10% of the course grade. Students with backgrounds in the area and specific expertise may wish to do in-class presentations for extra credit.

Class-By-Class Plan and Course Readings

The greyed out portion of this schedule is tentative and will be adjusted as the semester goes on. I may adjust assignments, assignment dates, and lecture topics based on the class's interests. Due dates for an assignment will be the date that the next assignment is handed out.

		Required Reading			

Date	Activity	(Section numbers refer to 2015 edition of Barber)	Optional Reading	Lecture Notes	Assignments
Aug 25	introductory meeting	Appendix A.1-A.4, 13.1-13.4	Chater, Tenenbaum, & Yuille (2006)	lecture	
Aug 27	basic probability, Bayes rule	1.1-1.5, 10.1	Griffiths & Yuille (2006)	lecture	Assignment 0
Sep 1	continuous distributions	8.1-8.3		lecture	
Sep 3	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
Sep 8	Gaussians	8.4-8.5		lecture	
Sep 10	motion illusions as optimal percepts	Weiss, Simoncelli, Adelson (2002)	motion demo 1 motion demo 2	lecture	Assignment 2
Sep 15	Bayesian statistics (conjugate priors, hierarchical Bayes)	9.1	useful reference: Murphy (2007)	lecture	
Sep 17	Bayes nets: Representation	2.1-2.3, 3.1-3.5	Cowell (1999) Jordan & Weiss (2002) 4.1-4.6	lecture	Assignment 3
Sep 22	Bayes nets: Exact Inference	5.1-5.5	Huang & Darwiche (1994)	lecture	
Sep 24					Assignment 4

Sep 29	Bayes nets: Approximate inference	27.1-27.6	Andrieu et al. (2003)	lecture	
Oct 1					
Oct 6	<catch up day>				Assignment 5
Oct 8	Learning I: Parameter learning	8.6, 9.2-9.4	Heckerman (1995) 9.5	lecture	
Oct 13	Learning II: Missing data, latent variables, EM, GMM	11.1-4, 20.1-3		lecture	
Oct 15	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	Assignment 6
Oct 20	text mining variational methods GUEST LECTURER: Jordan Boyd-Graber	28.1-28.5, 11.5	28.6-28.9	lecture1 lecture2	
Oct 22					
Oct 27	text mining topic model extensions	McCallum, Corrado-Emmanuel, & Wang (2005)	Bamman, Underwood, & Smith (2014)	lecture	
Oct 29	text mining nonparametric Bayes hierarchical models	Orbanz & Teh (2010) Teh (2006)		lecture1 lecture2	
Nov 3					Assignment 7

Nov 5	modeling and optimization Gaussian processes	19.1-19.5		lecture1 lecture2	
Nov 10	modeling and optimization Multiarm bandits and Bayesian optimization		Shahriari, Swersky, Wang, Adams, and de Freitas	lecture	
Nov 12	modeling and optimization\ Guest speakers: Mohammad Khajah, Manjhunath Ravi				Assignment 8
Nov 17	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	
Nov 19	sequential models exact and approximate inference (particle filters, changepoint detection)	27.6 Adams & MacKay (2008) Yu & Cohen (2009)	Wilder, Jones, & Mozer (2010)	ppt pdf	
Dec 1	sequential models Kalman filters	24.1-24.4	Koerding, Tenenbaum, & Shadmehr (2007) 24.5	lecture	
Dec 3	vision/attention search	Mozer & Baldwin (2008) Najemnik & Geisler (2005)	supplemental material for Najemnik & Geisler	lecture lecture	Assignment 9
	CLASS				

Dec 8	CANCELLED (or guest lecture)				
Dec 10	CLASS CANCELLED (or guest lecture)				
Dec 14 13:30-16:00	Final project presentations				Assignment 9 due

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
Murphy's probabilistic modeling toolbox
BUGS
OpenBayes
Orange
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)

Additional information for students (click to read)