

Syllabus

Advanced Machine Learning

CSCI 6622

Tue, Th 14:00-15:15
ECST 1B21

Instructor

[Professor Michael Mozer \(mozer@cs.colorado.edu\)](mailto:mozer@cs.colorado.edu)

Department of Computer Science

Engineering Center Office Tower 7-41

(303) 492-4103

Office Hours: Tu, Th 12:30-13:30

Course Objectives

This course aims to provide research experience in state-of-the-art statistical machine learning techniques to advanced graduate students in computer science and engineering. This experience will allow students to evaluate the field's potential and limitations first hand. The focus of the course will be semester-long student projects. Projects will focus on applications of machine learning techniques to real-world engineering and AI problems.

Prerequisites

This course is a continuation of last semester's CSCI 5622, Machine Learning. The first semester is a prerequisite for the second. However, you may enroll in 6622 without 5622 if you have a background in machine learning. In this case, you are expected to catch up on any material covered in 5622 that we will assume in 6622.

Course Requirements

Research Project

The project should be publishable, state-of-the-art research in the field. My expectation is that the final product of the course will be a journal- or conference-quality paper. Minimally, you are expected to produce a paper that could be submitted to one of the leading machine learning conferences (e.g., ICML, NIPS, UAI). You will see examples of papers that have appeared in these conferences over the course of the semester.

In past years, I've encouraged each student to come up with a project of their own choosing. I've had mixed experiences with this approach, but on the whole I've been disappointed with the outcomes. Some students have difficulty defining a project of appropriate scope, others become enamored with developing a novel algorithm for which there is no real need, and others don't make enough headway to have produced a substantive contribution by the semester's end. As a result, I'm going to try a new approach in 2006.

I would like projects to be *small group* efforts, with 2-4 students involved. I would like to focus on *applications* of machine learning to a substantive problem domain, rather than development of novel algorithms. And by default, I would like to propose problem domains for which we have significant data sets and clients who would like our help. In a small number of cases, students may wish their project to tie in with current research in your primary research area; we'll discuss those cases in class.

Students must take an immediate and active role in selecting a project, defining the scope of the project, and identifying and reading the relevant background literature. The standard sources for machine learning are: proceedings of the NIPS, ICML, or UAI conferences, and papers in the *Journal of Machine Learning Research* (JMLR) and *Machine Learning Journal* (MLJ).

The project will require that you:

- pick a topic of appropriate scale, not too trivial and not too ambitious
- search the relevant background literature to understand what ideas others have had
- situate your ideas and approach with respect to past work
- conduct experiments and evaluate your results, which may involve a comparison of multiple models and replicating the work of others
- write up your work in a scientific style appropriate for other researchers in the field.

Students are expected to develop their own software tools, or to find appropriate packages on the web. Matlab is the language of choice in machine learning, due to the large number of free packages available written in matlab, and its facilities for visualization.

Students will be responsible for presenting an initial proposal to the class and periodic updates throughout the semester. At the end of the semester, students will critique each other's project reports and will hold a mini-conference in which 15 minute summaries of each project are presented. (Fifteen minutes may not seem like much, but that's all you get at most professional conferences!) Local experts in machine learning may attend the mini-conference.

Readings

About half of class time will be spent on project proposals and updates. The other half will be used to discuss and evaluate papers from the current research literature. Readings (as they are assigned) are available [here](#).

Everyone is expected to read the papers in advance and to submit a one-page commentary on the papers (see below). We will take turns leading discussions of the papers. The leader is responsible for preparing a short summary of the work and for organizing the discussion. Paper topics can be dictated by student interest and it may be useful to select papers on topics related to one's project. A sample of possible topics: support-vector machines, Bayesian approaches, Gaussian processes, topic discovery, low-level vision, natural language understanding, speech recognition, reinforcement learning and control, models of human information processing and cognition.

Written Commentaries

The commentary consists of approximately one page of comments, questions, or critiques of the assigned reading(s) for that class. This page will be due the day of class, and can include one or more of the following:

- a summary of what you think the main or most interesting ideas are behind the reading(s).
- questions about the material for further discussion, either clarification questions or points of disagreement with the authors ("I don't see how such and such will work as the author claims...").
- comments on how the assigned reading relates to other course readings, or, if you feel ambitious and want to track down some related work in the field, how the assigned reading compares to this other work.
- a critique of the work. What are the flaws in the ideas presented? What are the limitations? Do the authors place their work in the appropriate theoretical perspective? Do the authors overstate their results? In what direction might the work be extended?

These commentaries are intended to promote careful thought about a paper *before* the session in which it is discussed. The point is not to give you more busy work, but rather to encourage you to jot down notes and questions as you read the papers. *They will not be accepted after the class in which the paper is discussed.*

Semester Grades

Grades will be based roughly on the following: written project 50%, oral progress reports and end-of-the-semester project summary 10%, oral discussions of papers from the literature 15%, written commentary on papers 25%.

Project Milestones

I outline here the major milestones in the project development. I assure you that if you fall behind this schedule, you will have serious trouble completing your project. It is not the type of assignment you can churn out in a week or two, if for no other reason than the simulations will probably require several weeks to complete. *I will not give incompletes*, so be forewarned.

WEEK	MILESTONE
1-3	discussion of possible research projects, including student-proposed projects; background reading
4	two-page project proposal due by the end of the week
5, 6	define scope of project and prepare training/testing data and determine method of evaluation
7, 8	complete implementation and initial simulations (possibly replicating others' work)
9, 10	wrap up simulations
11	spring break
12, 13	prepare preliminary write up (while continuing to run simulations if necessary)
14	exchange and critique one another's papers
15	prepare final version of paper, and oral presentations

Class-By-Class Plan and Course Readings

For the first couple of weeks, we'll determine the schedule class-by-class. I want to spend as much time as is necessary to make sure that everyone has formulated a project, and that participants have an understanding of each others' projects.

Date	Activity
Jan 17	Introduction
Jan 19	Introduction to graphical models (Kevin Murphy)
Jan 24	project presentations: Dr. Nicolas Nicolov, Chief Scientist, Umbria; Dr. Franco Salvetti
Jan 26	project presentations: On line phase detection algorithms (Nagpurkar et al.)
Jan 31	project presentations: Dr Sarel van Vuuren, CSLR
Feb 2	project presentations: lightening data
Feb 7	Maximum entropy spam filtering (Zhang & Yao, 2003) background reading on Maximum entropy models
Feb 9	Hidden Markov Models (Rabiner 1989)
Feb 14	Kalman Filters (Maybeck, 1979)
Feb 16	brief presentations on project-related research papers
Feb 21	brief presentations on project-related research papers
Feb 23	brief presentations on project-related research papers
Feb 28	Ensembles (Brown, Wyatt, Harris, & Yao, 2004) further background: Schapire 2002 ; Dietterich 2002
Mar 2	Factorial switching kalman filters for condition monitoring in neonatal intensive care (WilliamsQuinnMcIntosh2006)
Mar 7	Topic model GriffithsSteyvers2003
Mar 9	project proposal discussion
Mar 14	project proposal discussion discussion of LSA unbalanced data (Provost, 2000) -- NO NEED TO WRITE UP SUMMARY [optional: more on unbalanced data -- Japkowicz, 2000 , Cohen et al., 2000]
Mar 16	visit with Jeremy Kubica MEET IN ECOT 831
Mar 21	Gaussian Processes (read chapter 1 of Rasmussen & Williams, and sections 2.1 and 2.2 of chapter 2)
Mar 23	visit with Gary Cottrell from UCSD MEET IN ECOT 831
Mar 28, 30	SPRING BREAK
Apr 4	project progress reports (initial simulation results)
Apr 6	conditional random fields (Lafferty, McCallum, & Pereira, 2001)
Apr 11	NO CLASS -- MIKE IS AT CONFERENCE
Apr 13	Harmonium Model (Welling, Rosen-Zvi, & Hinton, 2005)
Apr 18	project progress reports
Apr 20	project progress reports / discussion of final paper
Apr 25	probabilistic interpretation of SVM outputs (Platt, 1999)
Apr 27	HMMs with Dirichlet mixture priors (Brown et al., 1993) optional: Mike's notes on Dirichlet distributions optional: Hierarchical Bayesian Markovian Model (Xing et al., 2003)
May 2	
May 4	
May 6, 4:30-7:00 p.m.	[final exam period] FINAL PRESENTATIONS

Queue

These are the papers I have in the queue. We will adjust the queue depending on what seems most pertinent to projects. IF YOU HAVE A PAPER YOU'D LIKE TO INCLUDE IN THE QUEUE, LET ME KNOW.

[Kersten et al. "Object perception as Bayesian Inference"](#)

[Gaussian Process Dynamical Models](#)

K. Murphy, "Switching Kalman Filters"

Zhang, Ghahramani, & Yang, "Multitask learning"

[Williams & Barber \(1997\)](#) Gaussian processes for classification

[Krishnapuram 2004 Bayesian feature selection](#)

[Scerbo \(2005\) Adaptive Automation](#)

REQUESTS

Random forests

Neural nets

novel applications: [LiaoFoxKautz2004](#)

Unsupervised learning: [HintonNair2006](#) (Inferring motor programs from images of handwritten digits)

time series prediction: [ThiessonChickeringHeckermanMeeks2004](#) (time series prediction with graphical models)

Reinforcement learning: SallansHinton2004 (RL with factored states and actions)

Neurobiology and machine learning

Other possible papers (my favorites from 2005):

[BengioDucharmeVincentJauvin2003](#)

[GruhlGuhaLibenNowellTomkins2004](#)

[SaulRoweis2001](#)

[HamLeeSaul2003](#)

[BleiNgJordan2002](#)

[BleiGriffithsJordanTenenbaum2004](#)

[NgJordanWeiss2002](#)

[Dietterich \(2002\) ensembles](#)

[Tipping \(2000\)](#) relevance vector machine (longer paper: [Tipping 2001](#))

[Tobler, Fiorillo, & Schultz \(2005\)](#)

Interesting Links

[Tutorials](#)

[background paper on statistical pattern recognition](#)

[Andrew Moore's machine learning tutorial slides](#)