CSE 515T: Bayesian Methods in Machine Learning (Spring 2015)

Instructor	Professor Roman Garnett
TA	Wenlin Chen
Time/Location	Monday/Wednesday 4–5:30pm, Cupples II 230
Office Hours (Garnett)	Thursdays 3–5pm, Jolley 504
URL	`http://cse.wustl.edu/~garnett/cse515t/`
GitHub	`https://github.com/rmgarnett/cse515t/`
Piazza message board	`https://piazza.com/wustl/spring2015/cse515t/home/`

Course Description

This course will cover modern machine learning techniques from a Bayesian probabilistic perspective. Bayesian probability allows us to model and reason about all types of uncertainty. The result is a powerful, consistent framework for approaching many problems that arise in machine learning, including parameter estimation, model comparison, and decision making. We will begin with a high-level introduction to Bayesian inference, then proceed to cover more-advanced topics.

Prerequisites

We will make heavy use of mathematics in this course. You should have a good grasp of multivariable calculus (integration, partial derivation, maximization, etc.), basic probability (conditional probability, expectations, etc.), and linear algebra (solving linear systems, eigendecompositions, etc.).

Please note that this is not an introduction to machine learning; the CSE 417A/517A courses fill that role. I will assume prior familiarity with the main concepts of machine learning: supervised and unsupervised learning, classification, regression, clustering, etc. The good news is that if you're not already familiar with these concepts, there are many free resources available to get caught up. I will list some of these on the course webpage.

Book

There is no required book. For each lecture, I will provide a list of related materials, including book chapters, videos, papers, code, etc. on the course webpage. These are to give you different viewpoints on the subject. Hopefully you can find one that suits you.

Although no book will be required, the following books are highly aligned with this course:

- *Pattern Recognition and Machine Learning* by Christopher M. Bishop. Covers many machine-learning topics thoroughly. Very Bayesian. Can also be very mathematical and take some effort to read.
- *Bayesian Reasoning and Machine Learning* by David Barber. Geared (as much as a machine-learning book could be) towards computer scientists. Lots of material on graphical models. Freely available online.[1]
- *Gaussian Processes for Machine Learning* by Carl Rasmussen and Christopher Williams. Excellent reference for Gaussian processes. Freely available online.[2]

[1] `http://www.cs.ucl.ac.uk/staff/d.barber/brml/`, link also on course webpage.
[2] `http://www.gaussianprocess.org/gpml/`, link also on course webpage.

The following books are good resources for Bayesian statistics:

- *Statistical Decision Theory and Bayesian Analysis* by James Berger. An old book (1980, advertises "with 23 illustrations" on the title page), but nonetheless an excellent introduction to Bayesian methods. Very clear. Probably provides the most convincing philosophical arguments for the Bayesian viewpoint I have ever read.
- *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation* by Christian Robert. Another fairly technical resource with passionate arguments for the Bayesian perspective.

Assignments

There will be five assignments throughout the semester, with two weeks available to complete each one. The lowest grade of these will be dropped. You can also use this as a "get out of homework free card" if you'd like. Use it wisely.

The assignments will form 30% of your grade, and each will have two types of questions: traditional "pencil-and-paper" questions, and programming exercises meant to give more insight into applying the techniques we will discuss on actual data. The former *will not be corrected.* If you make a reasonable attempt to answer a question, I will give you full credit. After each assignment, I will provide solutions online.

The programming exercises will require you to implement some of the theoretical ideas we discuss in class. The point of these exercises is both to lead to a better understanding by forcing a different viewpoint (that of the designer), and also to enable interaction. I encourage you to play with the data, parameters, etc. associated with these exercises to see how the results change. The point of the exercises is *not* for me to judge your programming skills, so *please do not hand in your code.* Rather, you should convey your answers via plots, tables, and/or discussion, as appropriate. As I don't need to read your code, feel free to use any language you'd like, but note that if I provide you with my own code, I will do so in MATLAB.

The assignment schedule is as follows:

#	available	due
1	January 14	January 28
2	February 2	February 16
3	February 18	March 4
4	March 4	March 25
5	March 30	April 13

Late policy

Assignments will be due during class on the dates above. I will allow you to turn in your assignment up to one class late with no penalty. After that, you'll have to use your dropped assignment.

Collaboration policy

Please feel free to collaborate on the paper-and-pencil questions! This is a good way to gain a deeper understanding of the material. Of course, you will be expected to write up your answers separately. Also feel free to collaborate on a high level on the programming exercises, but please write your own code and produce your own results.

Midterm

There will be a midterm held in class on Monday, 23 February. This will count for 30% of your grade. The questions on the midterm will be highly correlated with those on the previous assignments.

Project

In the second half of the semester, you will complete a project, which will comprise 30% of your final grade. The goal of the project will be to apply Bayesian techniques to a real dataset in a nontrivial way. I will compile a list of datasets on the course webpage, but you should of course feel free to find your own. The project should reach beyond the scope of the homework problems. I will judge the success of a project based on the methodological approach rather than the quantitative details of the final outcome. This is an exercise in applying theoretical ideas in practice, and even the most carefully constructed models or techniques can fail on a particular problem. Note that I would expect you to think about *why* your method might have failed (or succeeded!).

You can complete this project in groups of one, two, or three people. Of course, I will expect more out of larger groups.

There will be four components to this project:

- A project proposal, due **Friday, 6 March.** This should be an approximately one page document describing your idea. I will read this and give feedback/suggestions.
- A status report, due **Friday, 3 April.** I expect this to be one or two pages, updating me on the progress of your project, including data processing, implementation, experimental design decisions, etc.
- A 10-minute presentation describing the project. These will be held in class during the final week, on **Monday, 20 April** and **Wednesday, 22 April.** The presentation should briefly explain the idea, the data, and the results of your investigation.
- A final report, due **Friday, 1 May.** This should be an approximately four-page document explaining the idea, experimental setup, results, and your interpretation of them.

Grading

Your final grade will consist of the following weighted components:

component	%
assignments (lowest dropped)	30%
midterm	30%
project proposal	10%
project status report	10%
project presentation	10%
project final report	10%
final project total	40%

Topics

An outline of the topics I expect to cover is below; this is subject to change, more likely by deletion than addition. If there is a particular topic you would like me to spend more time on (or don't care about at all!), please let me know.

I will keep the course webpage updated with lecture-specific information and resources.

- **Introduction to the Bayesian method:** review of probability, Bayes' theorem, Bayesian inference, Bayesian parameter estimation, Bayesian decision theory, Bayesian model selection.
- **Approximate inference:** the Laplace approximation, variational Bayes, expectation propagation.
- **Sampling methods:** rejection sampling, importance sampling, Markov chain Monte Carlo.
- **Parametric models:** Bayesian linear regression, logistic regression, general linear models, basis expansions, mixture models, latent Dirichlet allocation.
- **Nonparametric models:** Gaussian proesses for regression and classification.
- **Bayesian numerical analysis:** Bayesian optimization, Bayesian quadrature.