CSE 5015: Bayesian Methods in Machine Learning (Fall 2025)

Instructor	Professor Roman Garnett
TA	Ryan Zhang
Time/Location	Monday/Wednesday 10–11:20pm, Seigle 301
Office Hours (Garnett)	by appointment, always available on Slack
Office Hours (TA)	TBA
URL	`https://www.cse.wustl.edu/~garnett/cse515t/fall_2025/`
GitHub	`https://github.com/rmgarnett/cse515t`
Slack	`http://bit.ly/3UFNnfn`

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.

This course is meant to lay the groundwork for research in these areas. If you are looking for a practical introduction with a focus on implementation, etc. this may not be the best course for you.

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.), 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 417T/517A and ESE 417 courses fill that role. I will assume prior familiarity with the main concepts of machine learning: supervised and unsupervised learning, classification, regression, clustering, etc.

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.

I have also posted links on the course webpage to several related books you may find useful, many of which are available for free online.

Reading

Regardless of what book(s) or other material(s) you may choose to consult, **you are expected to spend some time outside of class engaging in reading and self-study.**

Assignments

There will be a small number of assignments throughout the semester, with two weeks available to complete each one.

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.

Late policy

Assignments will be due during class on the dates specified on the course homepage. I will allow you to turn in your assignment up to one class late with no penalty.

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.

AI policy

If you would like to use ChatGPT or a similar resource while completing your assignments, that's fine with me. Just make sure you trust its output!

Midterm

There will be an in-class midterm on a date to be determined later (probably just before or just after fall break). This will count for 30% of your grade.

Project

In the second half of the semester, you will complete a project, which will comprise 40% of your final grade. You will have two possible paths to satisfy this project requirement – a free-form project or a supervised, guided project.

The main goal of the project is to give you hands-on experience applying Bayesian methods to a real-world dataset. The use of real-world data can have many interesting (and potentially frustrating!) aspects that are difficult to convey without getting your hands dirty. The scope of the project is intended to be more than a homework problem, but less than a full-fledged research paper.

We will talk more about the project a bit later in the semester.

Grading

Your final grade will consist of the following weighted components:

component	%
assignments	30%
midterm	30%
final project	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.

Other resources

The University has also put together a list of important policies and resources that I recommend you review if you haven't already.[1]

[1] https://provost.wustl.edu/syllabi-resources-and-template-language-danforth-campus/