DSA 595 Bayesian computations for machine learning

Semester: Spring 2026

Lecture days/times: Tuesdays 18:00–18:50

Location: 5270 SAS Hall

1 credit hour

Instructor: Dr. Jonathan P Williams
Email: jwilli27@ncsu.edu
Office location: 5218 SAS Hall
Office hours: 17:00–18:00 Tuesdays or by appointment
Office phone: 919.513.0191

Course Description: Build your background knowledge into your machine learning tasks with this primer on Bayesian statistical reasoning and computational tools. Students will be introduced to basic Bayesian principles including a collection of standard models. From a Bayesian perspective, students' data investigation questions of interest will motivate computations based on Markov chain Monte Carlo (MCMC) methods, including the Metropolis-Hastings and Gibbs MCMC algorithms. Students will have the opportunity to write their own computer code to implement these algorithms, train a variety of standard Bayesian models, and will ultimately create a custom course project implementing Bayesian computations for a machine learning task. No prior exposure to Bayesian inference is assumed.

Learning Outcomes. By the end of the course, students will be able to:

1. Implement Bayesian logic and arithmetic, e.g., derive a posterior distribution from a likelihood and prior distribution.
2. Write a Metropolis-Hastings algorithm in custom computer code.
3. Write a Gibbs sampling algorithm in custom computer code.
4. Diagnose convergence of an MCMC algorithm via trace plots and criteria.
5. Validate a Bayesian model in a simulation study based on an MCMC algorithm.
6. Train a Bayesian model on a real data set by implementing an MCMC algorithm.

Prerequisites: Basic proficiency in R, Python, or another interpretable programming language (e.g., know how to write a simple function), concepts from introductory statistics such as probability distributions, and measures of location and variability.

Optional Text: C. P. Robert and G. Casella (2004). *Monte Carlo Statistical Methods*, 2nd edition. Springer.

Digital Course Components:
Course website: `https://jonathanpw.github.io/DSA595`

Moodle (accessed via `https://wolfware.ncsu.edu`)

Grade Distribution:

Attendance	30%
Assignments	40%
Course project	30%

Letter Grade Distribution:

≥ 93.00	A	73.00 - 76.99	C
90.00 - 92.99	A-	70.00 - 72.99	C-
87.00 - 89.99	B+	67.00 - 69.99	D+
83.00 - 86.99	B	63.00 - 66.99	D
80.00 - 82.99	B-	60.00 - 62.99	D-
77.00 - 79.99	C+	≤ 59.99	F

For students taking the course as credit-only, S is equivalent to C- or better; otherwise U.
No expectations beyond attendance apply to students choosing to audit the course.

Final project due: Tuesday, 28 April, 2026

Personal note to students: Please do not feel intimidated about interacting with me. Regardless of how busy or stressed I may appear to you, teaching your class is a part of my job, and I take that very seriously. I care deeply about the quality of your learning. Please always reach out to me if you have questions, concerns, or need help. I understand that it can be difficult and can even feel embarrassing to ask for help. However, I was once in your position, and I promise to always treat you with respect, empathy, and kindness. Nobody that ever did anything meaningful did so without first failing over and over again.

Course policies and commentary:

- **Assignments**
  - Problem sets will be assigned every week, and the last 20 minutes of each lecture will be allocated to working on the problem set for the week. Collaboration is encouraged. Problem sets are due by the end of the day, each Friday.
  - Each homework assignment will receive the same weight in the calculation of the final course grade (i.e., longer (shorter) assignments do not count for a larger (smaller) portion of the overall assignment course grade).
- **Projects** Each student will propose a custom course project to implement some type of Bayesian computations for a machine learning problem of their own interest, and students are welcome to bring their own data. E.g., students may want to analyze a data set they have that is related to their job or dissertation research. I will help students formulate a Bayesian model based on their problem/data, and guide them through implementing the standard computational tools that we survey in this course, on their data.
- **Attendance**
  - Attendance is recorded. Please confirm your attendance is recorded by the instructor each lecture, beginning on the second lecture of the semester. Each lecture counts for

1 point of attendance, and the total course attendance grade is the average of points across all lectures, excluding the first lecture.

  - Use lecture time as you feel most productive, but do not use it in a way which is distracting to others.

- **Class recording statement**

  - No components of this course will be recorded.

Tentative Course Outline:

- Week 1: Refresher on probability and Bayes formula
- Week 2: Introduction to Bayesian inference: conjugate priors
- Week 3: Bayesian inference without conjugacy
- Week 4: Random number and variable generation
- Week 5: Metropolis-Hastings algorithm
- Week 6: Random walk Metropolis-Hastings
- Week 7: Convergence diagnostics
- Week 8: Multivariate posterior distributions and adaptive proposal distributions
- Week 9: Adaptive proposal distributions
- Week 10: Gibbs sampling
- Week 11: Metropolis-within-Gibbs
- Week 12: Monte Carlo integration
- Week 13: Importance sampling
- Week 14: TBD

NCSU Polices, Regulations, and Rules: Students are responsible for reviewing the NC State University Policies, Rules, and Regulations (PRRs) which pertain to their course rights and responsibilities, including those referenced both below and above in this syllabus:

- Equal Opportunity and Non-Discrimination Policy Statement `https://policies.ncsu.edu/policy/pol-04-25-05` with additional references at `https://equalopportunity.ncsu.edu/policies/`
- Code of Student Conduct `https://policies.ncsu.edu/policy/pol-11-35-01`
- Grades and Grade Point Average `https://policies.ncsu.edu/regulation/reg-02-50-03`
- Credit-Only Courses `https://policies.ncsu.edu/regulation/reg-02-20-15`
- Audits `https://policies.ncsu.edu/regulation/reg-02-20-04`

Policy on Academic Integrity: Cheating, plagiarism and other forms of academic dishonesty will not be tolerated. Violations of academic integrity will be handled in accordance with the Student Discipline Procedures (NCSU REG 11.35.02). Be aware of the Code of Student Conduct (NCSU POL11.35.01) and Pack Pledge.

Disability Services for Students: Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with the Disability Resource Office at Holmes Hall, Suite 304, 2751 Cates Avenue, Campus Box 7509, 919-515-7653. For more information on NC State's policy on working with students with disabilities, please see the Academic Accommodations for Students with Disabilities Regulation (NCSU REG 02.20.01).

Privacy: Students may be required to disclose personally identifiable information to other students in the course, via digital tools, such as email or web-postings, where relevant to the course. Examples include online discussions of class topics, and posting of student coursework. All students are expected to respect the privacy of each other by not sharing or using such information outside the course.