

Professor: Taylor R. Brown PhD
Live Sessions: Wednesday Evenings
Office Hours: Mondays Evenings
Email: trb5me@virginia.edu
Slack handle: Taylor

Slack: https://join.slack.com/t/ds6040/shared_invite/zt-18g9iomw5-x9PtxhjU3hfCDHm4hyTmHw (
Gradescope: www.gradescope.com (entry code J35273)
Anaconda: <https://www.anaconda.com/>
Collab: <https://collab.its.virginia.edu/portal>
Schedule: <https://docs.google.com/spreadsheets/d/12-XPgCC-DOKUsqHbvzYGVRvdUVD2xiN5Eumw>

Course Overview: Bayesian approaches explicitly account for the uncertainty present in most machine learning problems. This uncertainty derives from both randomness in observational processes and incompleteness in problem understanding. This course focuses on building models from data that provide Bayesian inferences and quantify the uncertainty in these inferences. Students will learn how to think probabilistically and apply this understanding to problems in a variety of areas.

Prerequisites: This course will be mathematical. You should have a good grasp of multivariate calculus and linear algebra. A previous course in statistics covering multiple linear regression, and familiarity with programming in Python is also required.

Learning Outcomes: Upon successful completion of this course, you will be able to:

1. Apply the appropriate probabilistic model to based on the characteristics of the problem.
2. Demonstrate the ability to convert actual data science problems from different domains into formal, mathematical representations.
3. Demonstrate the ability to apply appropriate analytical or computational solutions to obtain solutions to real problems .
4. Use results from current applied technical papers and video presentations, and understand the latest methods and the newest discoveries in Bayesian machine learning.

Important Links: Here are some useful links (textbook and code):

- <https://github.com/wbasener/BayesianML>
- <https://www.kaggle.com/billbasener/code>

Class Schedule: Live web-based sessions will be held on Wednesday evenings from 7:15–8:15pm and 8:30-9:30 EDT on Zoom. See the “Online Meetings” tab on Collab for information about how to connect. Specific dates are in Google sheets schedule that is linked to above. All official submission deadlines are provided via Gradescope.

Required Technical Resources and Technical Components: We will be running Python code in Jupyter Notebooks. I recommend that you download Anaconda (see above for link) because it bundles Python up with a bunch of third party packages and IDEs. In other words, it's a one-time download.

Evaluation Standards and Assessments: You will be scored on several categories of assignments: homeworks, projects, and participation. Assignments are to be submitted through Gradescope. **All deadlines are visible on Gradescope.**

- All five **homework** exercises provide problem-solving experiences that illustrate the concepts of Bayesian machine learning. Hence, they provide the opportunity to demonstrate understanding of class material. These assignments will include both analytical problem solving and data analysis, but I place more emphasis on the data analysis. For both parts, your answers should be written out, showing results with tables and graphs and explanations in the text. These assignments are worth 50% of your grade, and they occur in Modules 2-12. You will have about two weeks to complete each one of these assignments. One of the goals of this course is to learn how to present the results of analysis in a way that non-expert stakeholders can easily understand. To that end, it is not sufficient to simply provide code and plots, even if they are correct. Take time to interpret results (as is reasonable, you don't need to provide lengthy prose if the question asks for one number), and to write well commented code.
- **Participation** makes the course more interactive and enriching and includes module discussions, responses to peer posts, and attendance, participation, and answering questions in live sessions, as well as helping others in Slack. Your participation score will be worth 15% of your grade.
- The goal of the **project paper** is for you to apply Bayesian machine learning to a real dataset in an advanced way, and it counts for 20% of your final grade. The grading will verify you can apply probabilistic reasoning to a nontrivial problem of your own choosing. The paper will be graded based on the suitability of the chosen approach and the demonstration of your theoretical and technical understanding of your chosen method. It will not be graded on different performance metrics such as forecasting accuracy, computational efficiency, etc. You will work in groups of two to four people of your own choosing, but you may also choose to work alone (just be sure to talk to me about this, first). You will start by submitting a one-page project proposal describing the problem, data, and what you think you will use as your method of analysis. I will provide feedback on this. The final paper should be 4-6 pages. Please see Gradescope for more information about deadlines.
- You will also give a 5-7 minute recorded **project presentation**. This will be worth 15% of your final grade. Your presentation will be graded on how well it summarizes and communicates the findings from your written report, not on your theoretical understanding or computational achievements. All group members will receive the same project grades.

I will use the default letter grade thresholds to calculate your final grade: https://virginia.service-now.com/its/?id=itsweb_kb_article&sys_id=1153c16fdb41f444f32fb671d961934 A B- is the lowest satisfactory grade for graduate credit.

Communication and Student Response Time: Check Slack as often as you can and participate in as many discussions as you can. It is intended to be a place where you can reach out to each other and me to ask questions related to content and technology. Questions containing personal information can be emailed to me or sent to me via a direct message in Slack. Throughout our time together, the sooner you inform me of any problem that may affect your attendance or performance, the better the chance we have of solving it together.

Course Evaluations: Students may be expected to participate in an online mid-term evaluation. Students are expected to complete the online end-of-class evaluation. As the semester comes to a close, students will receive an email with instructions for completing this. Student feedback will be very valuable to the school, the instructor, and future students. We ask that all students please complete these evaluations in a timely manner. Please be assured that the information you submit online will be anonymous and kept confidential.

University of Virginia Honor System: All work should be pledged in the spirit of the honor system at the University of Virginia. The instructor will indicate which assignments and activities are to be done individually and which permit collaboration. The following pledge should be written out at the end of all quizzes, examinations, individual assignments, and papers: “I pledge that I have neither given nor received help on this examination (quiz, assignment, etc.).” The pledge must be signed by the student. For more information, visit www.virginia.edu/honor

Accommodations: It is my goal to create a learning experience that is as accessible as possible. If you anticipate any issues related to the format, materials, or requirements of this course, please meet with me outside of class so we can explore potential options. Students with disabilities may also wish to work with the Student Disability Access Center to discuss a range of options to removing barriers in this course, including official accommodations. Please visit their website for information on this process and to apply for services online: sdac.studenthealth.virginia.edu.

Optional Texts: Here are some supplementary texts.

- Barber, D., *Bayesian Reasoning and Machine Learning* (Cambridge: Cambridge University Press, 2012).
- Theodoridis, S., *Machine Learning: A Bayesian and Optimization Perspective* (Netherlands: Elsevier Science, 2015).
- Robert, C., *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation* (New York: Springer, 2007).
- Bishop, C.M., Bishop, P.o.N.C.C.M., *Pattern Recognition and Machine Learning* (Singapore: Springer, 2006).
- DeGroot, Morris, *Optimal Statistical Decisions* (New York: McGraw-Hill, 1970).
- Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin, *Bayesian Data Analysis, 3rd Ed.* (Boca Raton, FL: Chapman & Hall/CRC, 2011).
- Kruske, John K., *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (London: Academic Press, 2015).
- Martin, Osvaldo, *Bayesian Analysis with Python, 2nd Ed.* (Birmingham, UK: Packt, 2018).
- Ross, Sheldon, *Introduction to Probability Models, 9th Ed.* (Burlington, MA: Academic Press, 2007).