**DS6040 – Bayesian Machine Learning (Fall 2024)**

*At A Glance*

**Instructor:** Dr. Teague Henry (he/him/his)

**Email:** [trhenry@virginia.edu](mailto:trhenry@virginia.edu)

**TAs:** Navya Annapareddy ([na3au@virginia.edu](mailto:na3au@virginia.edu)), Lingzhen Zhu (lingzhen.zhu@virginia.edu)

**Office Hours:** Tuesday, 12:30-1:30pm (Data Science Building, First Floor Hub Space),

**Lectures Sessions:** Tuesdays/Thursdays, 11:00-12:15pm (Room 305)

**Important Deadlines:**

|  |  |
| --- | --- |
| **Assignment** | **Due Date** |
| Homework 1 | 09/20 |
| Homework 2 | 10/11 |
| Midterm | 10/22 |
| Homework 3 | 11/8 |
| Homework 4 | 12/06 |
| Final Due | 12/16 |
|  |  |

**Grade Breakdown:**

* Homework (80%) – 4 homework assignments, a mixture of coding and written responses.
* Midterm/Final (20%) – An in-class, hand written midterm, and a take home open book final.

**Assignment Submission:**

All assignments must be submitted electronically through Canvas by the specified due dates and times. It is crucial to complete all assigned work—failure to do so may result in failing the class.

**Late Assignments:**

Life happens, and there are many things more important than this class. As long as you have informed me (the instructor), **before the due date**, that a homework assignment will be late, that is fine, no points will be docked. The same goes for if an emergency disrupts your best laid plans to complete an assignment, just let me know.

***Late assignments that do not have a reason for being late will be docked one full letter grade per day late.***

All assignments, regardless of if a late submission was approved, must be submitted by the last day of classes, else they will be given a 0. I cannot fairly assess performance without all students turning in the same set of assignments.

**DS6040 – Bayesian Machine Learning (Fall 2024)**

***Syllabus***



**Instructor:** Dr. Teague Henry (he/him/his)

**Email:** [trhenry@virginia.edu](mailto:trhenry@virginia.edu)

**Class Time:** Tuesday/Thursday, 11am-12:15pm

**Canvas Site: Update!**

**Course Description:**

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 predictions and quantify the uncertainty in these predictions. Students will learn how to think probabilistically and apply this understanding to problems in a variety of areas.

**Prerequisites:**

This course will be heavily mathematical. A good grasp of multivariate calculus and linear algebra is required. A previous course in statistics covering multiple linear regression and programming with familiarity in both R and Python are also required.

**Required Texts (Assigned Chapters Available in the Canvas under Resources):**

* Basener & Brown, Bayesian Machine Learning ([Github Repo](https://github.com/UVADS/BayesianML))
* 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)

**Optional Readings:**

* 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, 201).3
* 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).

**Learning Outcomes:**

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

1. Apply the appropriate probabilistic technique to use 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. Report the results of data analysis or computation in a form that non-experts can easily understand.

**Grading Scale**

Your final letter grade will be determined by the following scale

|  |  |  |
| --- | --- | --- |
| **Final Grade** | **Lower Bound on % Grade** | **Upper Bound on % Grade** |
| A+ | 98.0 | 100.0 |
| A | 97.999 | 93.0 |
| A- | 92.999 | 90.0 |
| B+ | 89.999 | 87.0 |
| B | 86.999 | 83.0 |
| B- | 82.999 | 80.0 |
| C+ | 79.999 | 77.0 |
| C | 76.999 | 73.0 |
| C- | 72.999 | 70.0 |
| D+ | 69.999 | 67.0 |
| D | 66.999 | 63.0 |
| D- | 62.999 | 60.0 |

**Evaluation Standards and Assessments**

**Homework (80% of grade)** **–**

* There are four homework exercises spread across the modules to provide problem-solving experiences that illustrate the concepts of Bayesian machine learning.
* Homework is due 2 weeks from when it was assigned (at the end of said week).
* Each assignment will require programming in R or Python.
* 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.

**Midterm (10% of grade) –**

* This is an in-class, hand-written midterm. Please bring a writing utensil.
* You are allowed one piece of (typical) paper, back and front, for notes.
* If you have a calculator, you are welcome to bring it, but you will not be penalized if you leave calculations in fractional form. Note: The vast majority of the problems on this midterm will not require you to perform calculations. You may also use the calculator on your phone, just don’t use your phone for anything else!
* I reserve the right to curve the grade, but I will only ever increase your grade, never decreasing it. This is to calibrate the test to the difficulty.

**Final Exam (10% of the grade)**

* This is a take home “exam,” untimed with unlimited access to notes.
* The only thing that is not allowed for this exam is collaborating with class-mates, and, besides questions regarding technical difficulties, I and the TAs will not be answering questions about the content.
* Really, the only difference between this final exam and an assignment is that you’ll have a bit more of a limited window to work on it, and you can’t talk to your classmates about it.
* Like with the midterm, I reserve the right to curve this final exam, but I will only ever increase your grade.

**Teague’s Grading and Regrading Policy**

* Homework assignments are graded based on completion, not on correctness. If you make mistakes, we (myself or the TAs) will provide you with feedback so you can learn.
* The only way of losing points on a homework assignment is to not do a component of the assignment. If you are encountering difficulties with a part of the assignment, please each out to the TAs or myself. We are more than happy to answer questions about them.
* The Midterm and Final will be graded based on correctness. That being said, there are many ways the kind of problems I write for midterms/finals can be correct.
* The reasoning behind my grading policy is this: every person in the class should get 80% from their assignments, so that even if you literally do not do the midterm/final, you will still pass the class (as a B- is the lowest passing grade). In fact, you can get above 80% just from assignments because there are several extra-credit problems on these assignments.
* The midterm and final are there to let you distinguish yourself and flex your knowledge. They also provide me information as to what people are actually learning in the class, and I take performance into account when I re-prep these courses.

**Teague’s AI Policy**

The current generation of AI tools, like ChatGPT, are powerful tools that can be used to improve how you learn the content of this course. I encourage you to use these tools, with a couple of caveats:

* I consider the use of AI tools to perform *writing* to be a form of plagiarism. It also defeats the purpose of learning how to write (which is my main issue with it). Now, I am not running assignments through an AI detector, those don’t work. But, if you turn in something that is obviously AI generated, or something that just doesn’t match with the rest of the writing you’ve produced, I will give you a grade of 0 on that assignment, and start Honor Code proceedings. Don’t make me do that, I don’t want to, you don’t want me to, nobody benefits from plagiarism
* I consider the use of AI to analyze data (ala, throw a CSV file into chatGPT) to be a form of plagiarism (but more, if you do that why are you here? Like, do you want to learn how to do data science?).
* I **do not** consider the use of AI tools to perform editing and proofreading to be plagiarism, indeed, I use them to edit and proofread my own writing.
* I **do not** consider the use of AI tools to generate code to be plagiarism (or academic dishonesty). The simple fact of the matter is that if you don’t know what you are coding, generating the code will most likely result in code that doesn’t do what you want. I tend to use AI tools to write convenience functions that I know how to write, but don’t want to spend my time finicking with various technical details. So, if you use AI to write code, make sure it works. If you submit code that is obviously AI generated, and doesn’t work, I will consider that to be no-effort and grade the assignment down (but this is not an Honor violation).
* I absolutely **do not** consider the use of AI tools as a study tool to be any form of academic violation. I actually strongly encourage you to use it like it’s Wikipedia that you can have a back and forth conversation with. I’ve used it to familiarize myself with a number of topics that are directly related to my research.

In every case, I do ask you to tell me when you use AI tools, and how you’ve used them. This helps me understand better how students use them, and it’s just good practice.

**UVA Policies:**

**SDS Grading Policies**

The standing of a graduate student in each course is indicated by one of the following grades:  A+, A, A-; B+, B, B-; C+, C, C-; D+, D, D-; F. B- is the lowest satisfactory grade for graduate credit.

**Attendance**

Students are expected to attend all class sessions. Instructors establish attendance and   
participation requirements for each of their courses. Class requirements, regardless of delivery mode, are not waived due to a student's absence from class. Instructors will require students to make up any missed coursework and may deny credit to any student whose absences are excessive. Instructors must keep an attendance record for each student enrolled in the course to document attendance and participation in the class.

**University Email Policies**

Students are expected to check their ofﬁcial UVA email addresses on a frequent and consistent basis to remain informed of University communications, as certain communications may be time sensitive. Students who fail to check their email on a regular basis are responsible for any resulting  consequences.

**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](https://honor.virginia.edu/).

**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](https://www.studenthealth.virginia.edu/sdac). If you have already been approved for accommodations through SDAC, please send me your accommodation letter and meet with me so we can develop an implementation plan together.