STATISTC 540

Introduction to Statistical Learning

University of Massachusetts Amherst | Fall 2026

Instructor: David R. Burt (he/him)
Email: dburt@umass.edu
Office: LGRT 1545

Class Time: 1:00–2:15 p.m. Tu/Th, in person
Class Room: LGRT A201
Office hours: 2:30–3:30 p.m. Tu/Th LGRT 1522, By Appointment

Credits: 3

Course description

Introduction to modern statistical regression and classification techniques, including linear and logistic regression, nearest-neighbor methods, discriminant analysis, kernel smoothing, smoothing splines, local regression, generalized additive models, decision trees, random forests, and deep learning. Clustering methods such as k-means and hierarchical clustering will also be introduced. The course will also cover resampling-based methods for model evaluation and regularization-based methods for model selection. The course emphasizes both the mathematical foundations needed to understand the differences among these methods and their practical implementation.

Course Learning Objectives

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

- **Formulate statistical learning problems.** Translate real-world questions into regression, classification, or unsupervised learning tasks with clearly defined variables, objectives, and performance criteria.
- **Fit and interpret statistical learning models.** Implement linear and logistic regression, as well as other regression and classification methods in *R*. Explain how assumptions and model flexibility affect interpretation and performance.
- **Evaluate and select models.** Use appropriate performance measures; training-validation-test splits; and cross-validation to select between and evaluate models. Relate generalization performance to overfitting and the bias–variance tradeoff.
- **Communicate statistical ideas.** Explain statistical concepts, modeling choices, and empirical results clearly through reproducible written, visual, and oral presentations. State assumptions, uncertainty, and limitations.

Course Calendar

The course calendar, including exam dates, as well as what we plan to cover in each lecture, is available on the course website. Topics covered in lecture may change slightly if we end up ahead or behind schedule. I'll post any changes to the course calendar on the course website.

Prerequisites

STATISTC 416/516 or STATISTC 608 or permission of the instructor.

Textbook

The primary textbook for the class will be An Introduction to Statistical Learning with Applications in R (ISLR). We will also use The Elements of Statistical Learning as a reference. Both are available online for free:

- ISLR: https://www.statlearning.com
- ESL: https://hastie.su.domains/ElemStatLearn/

Software

Coding exercises will be written in R. Download and install R and RStudio. We will use Top Hat for attendance and in-class polling. You can access Top Hat at https://tophat.com/.

Learning Management System and Online Resources.

We have a Canvas page. The class notes, as well as useful links, will also be posted throughout the semester on my website at: https://davidrburt.github.io/statistc-540-fall-2026.html.

Additional Course Materials

The authors of ISLR have made freely available lecture videos for the content in the textbook. Watching these is always optional. We won't cover the exact same material in lectures (especially since these videos already exist!), but the videos are still a useful resource: https://www.youtube.com/playlist?list=PLoROMvodv4rOzrYsAxzQyHb8n_RWNuS1e. In class notes, I will suggest relevant sections of these lectures. These are a place where you can see similar material, described in a slightly different way. Especially if you are confused about how something was explained in class, they might be worth checking out! But also, if you are confused send me an email or come to office hours.

Office Hours

Please come to office hours! I'm happy to discuss the material, the lectures, or the course in general. If your question is specifically about a homework assignment, post it to the Canvas forum. You likely aren't the only student with that question, and this way everyone has access to any help on problem sets. Office hours are scheduled following class. If you have a question or concern and these times do not work in your schedule, send me an email to find another time to meet. If many of you are in this situation, I'll send out a poll and we will look to move one of the office hour blocks to a different time.

Contacting the Instructor by Email

Please email me with questions or concerns.

- Include the number of the course in your subject line (i.e. "STATISTC 540").
- I will do my best to respond within one working day.

Grading

Course grades will be based on the following components:

	Weight
Problem sets (homework)	20%
Three exams	30%
Course project	30%
Attendance	5%
Participation and engagement	15%

Final course percentages are converted to letter grades using the following ranges.

A	A-	B+	B	B-	C+	C	C-	D+	D	F
90–100	87–90	83–87	79–83	75–79	71–75	67–71	63–67	59–63	55–59	0–55

Graduate students are not eligible to receive grades of C-, D+, or D. Therefore, a graduate student with a final course percentage below 67 will receive an F.

Problem sets (Homework). Problem sets will be assigned approximately each week and submitted through Canvas. They will be posted at least one week before the due date. The lowest problem-set grade will be dropped. Late submissions will not ordinarily be accepted, because we will go over the problem sets in class, and accepting late submissions after this would create unequal access to the solutions. *A complete submission should include both the reproducible source file (an* `.Rmd` *or Quarto file) and its compiled PDF.*

Exams. There will be three equally weighted in-class exams based on the lecture notes and problem sets. Exam dates are October 1, October 27, and December 3.

Course project. The course will include a final project consisting of an in-class presentation and a written paper. Presentations will take place on December 8, 10, and 15. The written paper is due by 5:00 p.m. on Friday, December 18. A detailed description of the project requirements will be provided part way through the semester.

Attendance. We will generally take attendance during each class using Top Hat or a class roster. Attendance accounts for 5% of the final course grade, and each attendance point contributes one percentage point to the final course grade. The first unexcused absence does not reduce the attendance score; each additional unexcused absence reduces the score by one point, to a minimum of zero. If u is the number of unexcused absences, the number of attendance points is

$$A(u) = \max\{0,\, 5 - \max\{0,\, u - 1\}\}.$$

Participation and engagement. The goal of the class is for everyone to learn (about learning!). To that end, students are expected to participate and actively engage with the course material. The participation and engagement component accounts for 15% of the final course grade, and each participation point contributes one percentage point to the final course grade.

Students will receive one participation point for each unique lecture in which they ask or answer a substantive question about the course material, one point for each office-hours session they attend, one point for each Canvas forum in which they contribute, and one point for each typo they identify in the course materials. The total number of participation points will be capped at 15. If L, O, C and T denote the numbers of qualifying lectures, office-hours sessions, Canvas forums, and identified typos, respectively, the number of participation points is

$$P = \min\{15,\, L + O + C + T\}.$$

If you identify a typo, please email me with the course number in the subject line and a brief description of the typo and its location. If you ask or answer a question during a lecture, please email me with the course number in the subject line, the date of the lecture, and a brief description of the question. If you attend office hours, please sign in on the attendance sheet.

Excused Absences

In-person attendance and engagement are expected. Lectures will generally be recorded for students who must miss class because of extenuating circumstances and for review, but recordings are not a substitute for attending class.

Students who miss class because of a university-approved conflict—including athletic events, field trips, health reasons, religious observances, or other extenuating circumstances—should contact me promptly to make reasonable arrangements. Students remain responsible for meeting all course requirements.

Generative AI

Generative AI tools, including chatbots and coding assistants, are not permitted to be used on assignments in this class, unless an exception is explicitly stated on the assignment. In cases where an exception is stated, you will be required to submit an AI-use statement with your work, which will include what tools were used, how they were used, why you chose to use the tools you used, and checks you performed on the output of the tools. In all cases, you will be solely responsible for the content of your work. As with anything else, if you have questions regarding appropriate use of generative AI, or what constitutes use of generative AI, please ask!

University Policies

University policies regarding Accommodations, Academic Integrity, and Title IX, apply to all courses. These policies can be found in the https://www.umass.edu/senate/book/non-responsible-employee-required-syllabus-statements.