Syllabus
STA7934, Advanced Regression, Fall 2020
Aaron J. Molstad, University of Florida

Course objective: The course STA7934 will cover a variety of advanced methods used in modern regression analyses. The major topics covered in the course will be (in no particular order) semiparametric regression techniques (e.g., local linear regression, splines); regularization and shrinkage estimators (e.g., ridge regression and "lasso"); and advanced statistical machine learning methods (e.g., tree-based methods, support vector machines, neural networks). We may also cover graphical models, Bayesian additive regression trees (BART), and deep learning, if time permits. The course will culminate in a final project, which could include novel methodological research, an in-depth numerical or theoretical examination of a method covered in class, a novel data analysis project, or similar.

Lectures: Course lectures will take place on Zoom at the regularly scheduled class time (Monday, Wednesday, Friday, 12:50 - 1:40). A link to each lecture will be sent to the class at least 15 minutes before lecture begins over Canvas. These lectures will be recorded and posted to the course webpage. See **"virtual class consent"** below for more.

Office hours & appointments: Office hours will be held immediately after lecture from 1:45 - 2:45 on Monday and Wednesday of each week: the Zoom link provided for that day's lecture will also work for office hours. Office hours will not be recorded and can be made private if requested by the student. Appointments at times other than the dedicated office hours can be scheduled over email (`amolstad@ufl.edu`).

Prerequisites: A well prepared student should have taken advanced courses in generalized linear models, matrix algebra, theoretical statistics, and should be comfortable with the R programming language.

Course webpage: Course materials (e.g., homework assignments and due dates) will be posted to ajmolstad.github.io/F20_STA7934. A link to this page will be available on the course eLearning site, where students' grades will be posted. Please check this site regularly for updates and announcements. Course documents and important information, including homework exercises and solutions, will be posted here.

Schedule: Table 1 provides a tentative schedule of topics we will cover throughout the semester. This is subject to change, especially after week 10, depending on the pace of the course and students' interests.

Textbooks & other resources: There will be no single textbook used for the course. How-

Week	Topic
1	OLS, ridge regression
2	Lasso (computing, theory I)
3	Lasso (theory II, inference)
4	Discriminant analysis
5	Support vector machines, distance weighted discrimination
6	Piecewise polynomials and splines
7	RKHS, Gaussian process regression
8	Generalized additive models
9	Boosting, Random forests
10	Ensemble learning
11	Mixture models, convex clustering, hierarchical clustering
12	Gaussian graphical models and DAGs
13	Neural networks, deep learning
14	Deep learning continued, BART

Table 1: Tentative schedule of topics covered in this course.

ever, lectures will follow chapters from the following books:

- Hastie, Tibshirani, and Friedman, *Elements of Statistical Learning* [link]
- Ruppert, Wand, and Carroll, *Semiparametric Regression* [UF library]
- Wood, *Generalized Additive Models* [UF library]
- Hastie, Tibshirani, and Wainwright, *Statistical Learning with Sparsity* [link]

We will also cover material from recent articles: these will be cataloged on the aforementioned course page. Note that exams and homework will be based on material presented in lectures.

Assignments: The course will consist of four homework assignments, two exams, and a final project. The due dates are (tentatively):

- Homework 1: Friday, September 18th, 5:00pm (on Canvas)
- Homework 2: Friday, October 9th, 5:00pm (on Canvas)
- Exam 1: Friday, October 16th
- Homework 3: Friday, October 30th, 5:00pm (on Canvas)
- Exam 2: Friday, November 13th
- Homework 4: Friday, November 20th, 5:00pm (on Canvas)

- Presentation of projects: December 7th & 9th
- Submit project: December 14th, 5:00pm (on Canvas)

Some important notes about homework and exams:

- Homework assignments will often require a combination of programming (in R), statistical theory, and data analysis. Ideally, these will be completed and submitted as a Jupyter Notebook or Rmarkdown (.rmd) document. The instructor will give a brief overview of using Jupyter Notebooks in the first week of class.
- Homework assignments will be posted on the course webpage at least two weeks before their due date.
- *It is highly encouraged to work on these will classmates, although all submitted work should be yours alone.*
- Tentatively, the plan is to have take-home exams: more information will be provided as the first exam approaches.

Course project: The course project will be completed in lieu of a final exam. *Ideally, this project will complement your current research, or allow you to learn more about a topic from the course (or course-adjacent) which particularly interested you.* Projects can be methodological, numerical, or applied; e.g., you could apply a method from the course to a new dataset; you could perform extensive simulation studies comparing methods from the course, etc.

The project will consist of four checkpoints:

- *Project approval:* Monday, November 2nd

  You will writeup a 2-3 paragraph project proposal laying out *specifically* what you will do. These will be discussed in a one-on-one meeting with the instructor, who will approve the project or suggest changes.

- *Project check-in:* Monday, November 23rd

  You will writeup a 2-3 paragraph project progress statement laying out what specific progress has been made thus far. At this stage, you may also propose amendments to the original project proposal.

- *Project presentation:* Wednesday, December 9th

  You will prepare a 15 minute presentation describing your project, the progress you've made (and/or challenges you've faced), and directions for future research.

- *Submit project:* Monday, December 14th

You will submit a pdf writeup of your project. This should be a latex document in the style of a journal article. There is no page minimum or maximum per-se, but should be self-contained, have an introduction, methods, and results section, along with a bibliography.

Grades: Course grades will be based on four homework assignments (40%), two examinations (15% each), and the course project (30%). Grade cutoffs are provided in Table 2.

Grades	Percentage cutoff
A	100 - 90
A-	90 - 87
B+	87 - 84
B	84 - 80
⋮	⋮
D-	60 - 55
F	55 - 0

Table 2: Grade cutoffs.

Missed exam: Missed exams will receive a grade of zero, except in the cases of emergency. If an exam will be missed for a non-emergency reason, the student must notify the instructor *at least two weeks prior* to the exam date and provide proper documentation. Refer to for specifics on course attendance policy and missed exam policy.

Incomplete policy: The grade of "Incomplete" can be assigned at the discretion of the instructor when, due to extraordinary circumstances, e.g., hospitalization, a student is prevented from completing the work of the course on time. In particular, an "Incomplete" will be given if the student presents evidence from a certified professional that the student is unable to complete the course. In addition, the student must have a passing grade at the time of the incomplete request in order to receive the "Incomplete". An "Incomplete" requires a written agreement between instructor and student found here https://clas.ufl.edu/files/2019/02/CLASIncompleteGradeContract.pdf.

Students with Disabilities: Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Online course evaluation process: Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at https://evaluations.ufl.edu.

Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these evaluations are available to students at https://evaluations.ufl.edu/results/.

Virtual class consent Our class sessions may be audio-visually recorded for students in the class to refer back and for enrolled students who are unable to attend live. Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate verbally are agreeing to have their voices recorded.

If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live. The chat will not be recorded or shared.

As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.