

# STAT 4540: Statistical Learning (IGPI 4540)

Fall 2017

4:30 p.m. – 5:20 p.m. MWF in 140 Schaeffer Hall (SH)

<b>Instructor:</b>	Sanvesh Srivastava
<b>Email:</b>	sanvesh-srivastava@uiowa.edu
<b>Office:</b>	219 SH (Schaeffer Hall)
<b>Office hours:</b>	Mondays 3:00–4:00 p.m. & Wednesdays 2:00–4:00 p.m. in 219 SH
<b>Website:</b>	Canvas <a href="https://uiowa.instructure.com/">https://uiowa.instructure.com/</a>
<b>Lectures:</b>	4:30–5:20 p.m. MWF in 140 SH
<b>Lab:</b>	Labs will replace some lectures; tentatively, there will be a lab at the end of every chapter.

## Departmental Information and Executive Officer

**College:** College of Liberal Arts and Sciences

**Department:** Statistics and Actuarial Science

**Executive Officer:** Joseph Lang (joseph-lang@uiowa.edu)

**Address:** 241 SH

**Contact:** (319)-335-0712

**Email:** [statistics@uiowa.edu](mailto:statistics@uiowa.edu); [actuarial-science@uiowa.edu](mailto:actuarial-science@uiowa.edu)

**Website:** <http://www.stat.uiowa.edu>

## Textbook and course materials

*An Introduction to Statistical Learning, with applications in R* by James, Witten, Hastie, and Tibshirani. The website accompanying the book is <http://www-bcf.usc.edu/~gareth/ISL/>. This website is an excellent resource for many materials that we will use throughout this course, including a pdf copy of the book.

I will post announcements, homework problems, lecture notes, and other course information in Canvas.

## Course Description and Objectives

The course is an introduction to supervised and unsupervised statistical learning, with a focus on regression, classification, and clustering. Methods will be applied to real data using appropriate software. Supervised learning topics include: linear and non-linear (e.g. logistic) regression; linear discriminant analysis; cross-validation, bootstrapping, model selection, and regularization methods (e.g. ridge and lasso); generalized additive and spline models; tree-based methods, random

forests and boosting; and support-vector machines. Unsupervised learning topics include: principal components and clustering.

The students will develop an understanding of various statistical learning methods and will learn the computational skills to apply these methods to real-world data sets, to pursue a career in applied statistics, and to pursue research in statistical sciences and other data sciences.

It is expected that students will read the book, work on problems as required to master the material, and spend time applying the statistical learning methods to real-world data sets. You are expected to put in 6-8 hours of work outside of class. A few of you will do well with less time than this, and a few of you will need more.

### **Pre-requisites**

An introductory statistics course and a regression course. Prior exposure to programming and/or software, such as R, SAS, and Matlab is recommended, but not required.

### **Attendance**

Attending classes is required, but I won't take attendance.

### **Office hours**

Instructor office hours are every Mondays 3:00–4:00 p.m. & Wednesdays 2:00–4:00 p.m. in 219 SH. If you are unable to attend office hours during these times, then you should email the instructor for an appointment.

### **Grades**

Your grade will consist of Homework (20%), Project (15%), Quiz (10%), a Midterm Exam (25%) and a Final Exam (30%). At the end of the semester, Homework, Project, Quiz, Midterm, and Final grades may be normalized within each category in order to calculate the final course grade.

A *plus-minus* grading system will be used. Here is a *tentative grading scale*: A,A-: 88%–100%, B+,B,B-: 77%–88%, C+,C,C-: 61%–77%, D: 50%–61%, F:<50%. A+ will be given only in exceptional cases.

### **Homework**

Homework will be usually assigned every other week on Fridays and will be due two weeks later on Friday at the beginning of the class. Any exceptions will be announced in class or in Canvas. Your work must be legible, include name, and be stapled to receive credit. Due to time constraints, the grader may grade only some of the assigned questions, but you are responsible for understanding all questions.

All homeworks will contribute towards your final grade. Unless prior arrangements are made well in advance, for reasons judged to be acceptable by me, late homework will receive zero credit as solutions will be posted soon after the homework is due.

## Quizzes and Exams

We will (tentatively) have six announced quizzes in class. They will emphasize examples and key concepts repeatedly mentioned in class and in homework problems. There will be one midterm one final exam. The final exam will be comprehensive and cumulative. All exams will be closed book. You will be allowed *one* A4-sized handwritten “cheat-sheet” in the exams and quizzes.

## Project

The project in this class will replace a midterm exam. Students are expected to work in groups. The maximum size of any group is 5.

## Lab

Labs will be held depending on the needs of the students. Labs will be announced in class or in Canvas.

## Extra help

The *Statistics Tutorial Lab*, located in 202 CC. In addition, several graduate students have volunteered to independently tutor students in various courses at mutually-arranged times and fees. Please check the web site [www.stat.uiowa.edu/courses/tutoring.html](http://www.stat.uiowa.edu/courses/tutoring.html) for tutoring details.

## Grading Errors

Although every effort will be made to mark your work accurately, sometimes grading mistakes happen. If you believe that an error has been made on an in-class problem or exam, return the paper to the instructor immediately, stating your claim in writing.

## Important Dates

Check the office of the registrar website for the academic calendar:

<http://registrar.uiowa.edu/academic-calendar#!fall-2017>

Some important **tentative** dates for STAT 4540 are as follows:

<b>Midterm Exam:</b>	Wed, Oct 11
<b>Final Exam:</b>	Mon, Dec 11 – Fri, Dec 15 (to be decided by the university)

## Tentative Schedule

Chapters	Week	HW (Assigned)	Deadlines
1-2	08/21 – 08/25	1	
2-3	08/28 – 09/01		Project groups due on 09/01
3	09/04 – 09/08	2	HW 1 due on 09/08
3-4	09/11 – 09/15		Project topic and abstract due on 09/15
4	09/18 – 09/22	3	HW 2 due on 09/22
4	09/25 – 09/29		
5-6	10/02 – 10/06	4	HW 3 due on 10/06
6	10/09 – 10/13		Project proposals due on 10/09; Midterm Exam on 10/11
6	10/16 – 10/20	5	HW 4 due on 10/20
6	10/23 – 10/27		
7	10/30 – 11/03	6	HW 5 due on 11/03
7-10	11/06 – 11/10		Preliminary project reports due on 11/10
10	11/13 – 11/17	7	HW 6 due on 11/17
Break	11/20 – 11/24		–
10	11/27 – 12/01		Project reports due on 11/29;
–	12/04 – 12/08		Project presentations; HW 7 due on 12/06
	12/11 – 12/15		Final Exam

## College of Liberal Arts and Sciences: Policies and Procedures

<http://clas.uiowa.edu/faculty/teaching-policies-resources>

**Administrative Home of the Course** The College of Liberal Arts and Sciences is the administrative home of this course and governs matters such as the add/drop deadlines, the second-grade-only option, and other related issues. Different colleges may have different policies. Questions may be addressed in 120 SH or see the CLAS Student Academic Handbook:

[www.clas.uiowa.edu/students/academic\\_handbook/index.shtml](http://www.clas.uiowa.edu/students/academic_handbook/index.shtml)

**Electronic Communication** University policy specifies that students are responsible for all official correspondences sent to their standard University of Iowa email address (@uiowa.edu). Students should check this account frequently. (Operations Manual, III.15.2. Scroll down to k.11.)

**Academic Fraud** Plagiarism and any other activities when students present work that is not their own are academic fraud. Academic fraud is a serious matter and is reported to the departmental DEO and to the Associate Dean for Undergraduate Programs and Curriculum. Instructors and DEOs decide on appropriate consequences at the departmental level while the Associate Dean enforces additional consequences at the collegiate level. See the CLAS Student Academic Handbook.

**Making a Suggestion or a Complaint** Students with a suggestion or complaint should first visit the instructor, then the course supervisor, and then the departmental DEO. Complaints must be made within six months of the incident. See the CLAS Student Academic Handbook.

**Accommodations for Disabilities** A student seeking academic accommodations should first register with Student Disability Services and then meet privately with the course instructor to make particular arrangements. [www.uiowa.edu/~sds/](http://www.uiowa.edu/~sds/)

**Understanding Sexual Harassment** Sexual harassment subverts the mission of the University and threatens the well-being of students, faculty, and staff. All members of the UI community have a responsibility

to uphold this mission and to contribute to a safe environment that enhances learning. Incidents of sexual harassment should be reported immediately. See the UI Comprehensive Guide on Sexual Harassment at [www.uiowa.edu/~eod/policies/sexual-harassment-guide/index.html](http://www.uiowa.edu/~eod/policies/sexual-harassment-guide/index.html) for assistance, definitions, and the full University policy.

**Reacting Safely to Severe Weather** In severe weather, class members should seek appropriate shelter immediately, leaving the classroom if necessary. The class will continue if possible when the event is over. (Operations Manual, IV.16.14. Scroll down to e. h. and i.)