Syllabus: STAT 4620

Introduction to statistical learning

Autumn 2024

2 Credit Hours

## Course overview

### Instructor

* **Name:** Massimiliano (Max) Russo
* **Email Address:** russo.325@osu.edu
* **Office Hours:** In person at 317 Cockins Hall: Wednesdays, 3 pm-4 pm; or by appointment.

**Note**: My preferred method of contact is by email.

### Graduate Teaching Assistant

**Name:** Yingyu Cheng
**Email Address:** cheng.1753@osu.edu

### Course description

STAT 4620 provides an introduction to the principles of statistical learning and standard learning techniques for regression, classification, clustering, dimension reduction, and feature extraction. Specific topics that will be covered include an overview of predictive modeling and model evaluation, penalized regression and nonparametric regression, classification and regression trees, model selection and validation, and high-dimensional data and variable selection.

### PrerequisiteC- or better in STAT 3302: Statistical Modeling for Discovery II.Course expected learning outcomes

By the end of this course, students should successfully be able to:

1. Recognize the types of learning problems and understand their statistical formulations;
2. Understand the foundational principles of statistical learning, including statistical modeling, computation, and evaluation;
3. Comprehend the rationale and algorithms behind statistical learning techniques and know their relative merits and limitations;
4. Evaluate and compare different learning techniques numerically in terms of generalization error;
5. Use statistical learning methods for data analysis and interpret the results in the context of the data problem.

## Course materials and technologies

### Textbooks

James, G., Witten, D., Hastie, T., and Tibshirani, R. An Introduction to Statistical Learning with Applications in R. Springer

The textbook is available for free through SpringerLink:
<https://link.springer.com/book/10.1007/978-1-4614-7138-7>.

### Course technology

#### Technology support

For help with your password, university email, Carmen, or any other technology issues, questions, or requests, contact the Ohio State IT Service Desk. Standard support hours are available at [it.osu.edu/help](http://it.osu.edu/help), and support for urgent issues is available 24/7.

* Self-Service and Chat support: [it.osu.edu/help](http://it.osu.edu/help)
* Phone: 614-688-4357(HELP)
* Email: 8help@osu.edu
* TDD: 614-688-8743

#### Technology skills needed for this course

* Basic computer and web-browsing skills
* Navigating Carmen ([go.osu.edu/canvasstudent](http://go.osu.edu/canvasstudent))
* CarmenZoom virtual meetings ([go.osu.edu/zoom-meetings](http://go.osu.edu/zoom-meetings))

#### Required Equipment

* Computer: current Mac (MacOs) or PC (Windows 10) with high-speed internet connection
* Other: a mobile device (smartphone or tablet) to use for BuckeyePass authentication

#### Required software

#### This class requires you to use the free statistical software package R (The R Project for Statistical Computing; <http://www.r-project.org/>)

* You can download R for Windows, Mac, and Linux, from the CRAN archive at <https://cran.r-project.org>.
* An in-depth introduction to R is available at <http://cran.rproject.org/doc/manuals/R-intro.pdf>.
* Hands-on tutorials are available in the Swirl system, which you can learn about at <http://swirlstats.com/>. In particular, “R Programming: The basics of programming in R” is an appropriate first tutorial for students who have never used R.

An easier to use interface to R is available in the free software package RStudio. This package is available for Windows, Mac, and Linux and can be downloaded from [https://posit.co](https://posit.co/). **Note that RStudio requires R to be installed.**

### Course delivery

* This class will take place in person two times per week on **Wednesdays and Fridays, 1:50PM-2:45PM in Lazenby Hall 002**.
* All assignments will be posted on the Carmen class website. You will be given ample time to complete the assignments. Assignment due dates will be announced in class and on the Carmen course webpage.
* I will hold weekly in person office in my office in Cockins Hall 317.

## Grading and instructor response

### How your grade is calculated

| **Assignment Category** | **Points and/or Percentage** |
| --- | --- |
| Participation and discussion | 5% |
| Homework  | 15% |
| Exam 1  | 25% |
| Exam 2 | 25% |
| Final project | 30% |
| **Total** | **100%** |

All course grades will be recorded on the class website (Carmen)

### Assignment Information

#### Homework

I will assign homework (approximately) bi-weekly throughout the semester. You may discuss the homework with other students, but DO NOT copy any part of someone else's work or solutions from any other sources. Violations of this policy will be treated as academic misconduct. I encourage you to talk to me or the GTA (during office hours) if you have questions after serious attempts have been made to work on an assignment.

Please write clear and detailed answers to the homework problems and provide a statement interpreting the obtained results, if appropriate. You need to show your justification for or work on each homework problem. **Answers without work will not receive full credit**.

Exams

There will be two one-hour in-class exams, with tentative dates listed on the schedule at theend of the syllabus; any date changes will be communicated well in advance in class and on theCarmen webpage. The course project will serve as a cumulative evaluation of your learning inlieu of a final examination. During the exams, you will need to work independently without anyform of assistance or communication with anyone. A basic calculator is permitted; however,using a cell phone, tablet, laptop, or any other device for this purpose is not permitted.

Final Project

Students will be responsible for completing a team project. Each team will consist of 4-5 students. Details will be provided in class mid-way through the semester, and the project will be due near the end of the semester. The project will consist of formulating questions that can be answered with data and performing an appropriate analysis to answer the questions. Each team will be asked to submit a final report and give a brief introduction of their project in the last week of classes in a “rapid fire” presentation session. The overall project score will be based on the evaluation of the following components: formulation of relevant questions, appropriate analyses, and communication of findings (both in the “rapid fire” session and the project report)

### Use of GenAI tools in assignments

The Internet plays a significant role in learning, especially when dealing with coding and debugging. It is acceptable to use online resources, including generative artificial intelligence (GenAI), for tasks such as clarifying lecture material, seeking examples, and brainstorming ideas. Just as you would consult with a classmate for these purposes without violating academic integrity, you may also turn to GenAI. Students are permitted to use GenAI tools with this declination for most course assignments, **except for exams**.

**If you use GenAI in any of your assignments, please include a statement in your assignment** specifying the name of the tool you used (e.g. Copilot, ChatGPT, Grammarly, etc..) and how you used it. **You are responsible** for ensuring that the information you submit based on a GenAI query does not contain misinformation, unethical content, or violate intellectual property laws. Submission of GenAI-generated content as your own work is considered a violation of Ohio State’s [Academic Integrity](https://oaa.osu.edu/artificial-intelligence-and-academic-integrity) policy and [Code of Student Conduct](https://trustees.osu.edu/bylaws-and-rules/code) because the work is not your own. The use of unauthorized GenAI tools will result in referral to the [Committee on Academic Misconduct](https://oaa.osu.edu/academic-integrity-and-misconduct).

**If I suspect that you have used GenAI inappropriately on an assignment for this course, I will ask you to communicate with me to explain your process for completing the assignment in question.**

### Late assignments

Homework and project assignments will be submitted via the Carmen course webpage. Generally**, late assignments will not be accepted**, and no make-up exams will be provided. However, if there are extenuating circumstances beyond your control, please contact the course instructor immediately.

### Grading Scale

* 93-100: A
* 90-92: A$-$
* 87-89: B+
* 83-86: B
* 80-82: B$-$
* 77-79: C+
* 73-76: C
* 70-72: C$-$
* 67-69: D+
* 60-66: D
* Under 60: E

### Instructor feedback and response time

I am providing the following list to give you an idea of my intended availability throughout the course. (Remember that you can call 614-688-HELP at any time if you have a technical problem.)

Grading and feedback
For large weekly assignments, you can generally expect feedback within 7-10 days

E-mailI will typically reply to e-mails within **24 hours on school days**.

## Attendance and participation

###  Student participation requirements

## Your participation is based on your in-person attendance. The following is a summary of

## everyone's expected participation:

## Logging-in: AT LEAST ONCE PER WEEKBe sure you log in to the course in Carmen each week, including weeks with holidays. (During most weeks, you will probably log in many times.) If you have a situation that might cause you to miss an entire week of class, discuss it with me as soon as possible.

## In-person class meetings: REQUIREDYou are required to attend all in-person lectures, and you are responsible for all material presented during these lectures. However, formal attendance will not be taken during the class.

## Office hours: OPTIONALMy office hours will be held in person in my office in Cockins Hall 317. If you would like to discuss an assignment with me outside of my scheduled office hours, please contact me at the beginning of the week (as soon as possible).

## Course Schedule

Refer to our Carmen course page for up-to-date assignment due dates.

| Week | Date | Topics | Reading |
| --- | --- | --- | --- |
| 1 | 8/218/23 | Introduction; Linear Regression: SLR, MLR, Geometry, Loss, Weighted Least Squares | Chapters 1–3 |
| 2 | 8/28 8/30 | Beyond Linear Regression; Cross-Validation | Chapter 4, Sec 5.1 |
| 3 | 9/4 9/6 | Classification: Logistic Regression, Linear Discriminant Analysis (LDA) | Chapter 4 |
| 4 | 9/11 9/13 | Bootstrap, Cross-Validation & Bootstrap Examples | Chapter 5 |
| 5 | 9/18 9/20 | Regularization: Ridge Regression, LASSO | Chapter 6 |
| 6 | 9/25 | Regularization: Dimension Reduction, PCA, PCR | Chapter 6 |
| 6 | 9/27 | Exam 1 |  |
| 7 | 10/210/4 | High Dimensional Data Analysis and the Curse of Dimensionality; Spline Regression; Smoothing Splines; Local Regression  | Sec 6.4, Chapter 7 |
| 8 | 10/9 | Smoothing Splines; Local Regression  | Chapter 7 |
| 8 | 10/11 | Autumn Break – no class |  |
| 9 | 10/16 10/18 | Generalized Additive Models (GAM), Regression and Classification Trees  | Chapter 7,8 |
| 10 | 10/2310/25 | Trees: Boosting, Random Forests  | Chapter 8 |
| 11 | 10/3011/1 | Trees; Clustering: K-means, Hierarchical  | Chapter 8,10 |
| 12 | 11/6 | Clustering: K-means, Hierarchical | Chapter 10 |
| 12 | 11/8 | Exam 2 |  |
| 13 | 11/1311/15 | Clustering: K-means, Hierarchical | Chapter 10 |
| 14 | 11/20 | Additional topics – time permitting |  |
| 14 | 11/22 | Final project presentation |  |
| 13 | 11/2711/29 | Thanksgiving Holiday – no class  |  |
| 12 | 12/4 | Final project presentation |  |

## Academic policies

### Discussion and Communication Guidelines

The following are my expectations for how we should communicate as a class. Above all, please remember to be respectful and thoughtful.

* **Writing style:** While there is no need to participate in class discussions, as if you were writing a research paper, you should remember to write using good grammar, spelling, and punctuation. A more conversational tone is fine for non-academic topics.
* **Tone and civility:** Let's maintain a supportive learning community where everyone feels safe and where people can disagree amicably. Remember that sarcasm doesn't always come across online. I will provide specific guidance for discussions on controversial or personal topics.
* **Citing your sources:** When we have academic discussions, please cite your sources to back up what you say. For the textbook or other course materials, list at least the title and page numbers. For online sources, include a link.
* **Backing up your work:** Consider composing your academic posts in a word processor, where you can save your work, and then copying into the Carmen discussion.

### Academic integrity policy

See **Descriptions of major course assignments**, above, for my specific guidelines about collaboration and academic integrity in the context of this online class.

Academic integrity is essential to maintaining an environment that fosters excellence in teaching, research, and other educational and scholarly activities. Thus, The Ohio State University and the Committee on Academic Misconduct (COAM) expect that all students have read and understand the University's Code of Student Conduct, and that all students will complete all academic and scholarly assignments with fairness and honesty. Students must recognize that failure to follow the rules and guidelines established in the University's Code of Student Conduct and this syllabus may constitute Academic Misconduct.

The Ohio State University's Code of Student Conduct (Section 3335-23-04) defines academic misconduct as: Any activity that tends to compromise the academic integrity of the University or subvert the educational process. Examples of academic misconduct include (but are not limited to) plagiarism, collusion (unauthorized collaboration), copying the work of another student, and possession of unauthorized materials during an examination. Ignorance of the University's Code of Student Conduct is never considered an excuse for academic misconduct, so I recommend that you review the Code of Student Conduct and, specifically, the sections dealing with academic misconduct.

If I suspect that a student has committed academic misconduct in this course, I am obligated by University Rules to report my suspicions to the Committee on Academic Misconduct. If COAM determines that you have violated the University's Code of Student Conduct (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in this course and suspension or dismissal from the University.

If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact the instructor for this course.

* Committee on Academic Misconduct web page ([go.osu.edu/coam](http://go.osu.edu/coam))
* Ten Suggestions for Preserving Academic Integrity ([go.osu.edu/ten-suggestions](http://go.osu.edu/ten-suggestions))

### Copyright for instructional materials

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

### Creating an Environment Free from Harassment, Discrimination, and Sexual Misconduct

The Ohio State University is committed to building and maintaining a community to reflect diversity and to improve opportunities for all. All Buckeyes have the right to be free from harassment, discrimination, and sexual misconduct. Ohio State does not discriminate on the basis of age, ancestry, color, disability, ethnicity, gender, gender identity or expression, genetic information, HIV/AIDS status, military status, national origin, pregnancy (childbirth, false pregnancy, termination of pregnancy, or recovery therefrom), race, religion, sex, sexual orientation, or protected veteran status, or any other bases under the law, in its activities, academic programs, admission, and employment. Members of the university community also have the right to be free from all forms of sexual misconduct: sexual harassment, sexual assault, relationship violence, stalking, and sexual exploitation.

To report harassment, discrimination, sexual misconduct, or retaliation and/or seek confidential and non-confidential resources and supportive measures, contact the Office of Institutional Equity:

Online reporting form at [equity.osu.edu](http://equity.osu.edu/),

Call 614-247-5838 or TTY 614-688-8605,

Or Email equity@osu.edu

The university is committed to stopping sexual misconduct, preventing its recurrence, eliminating any hostile environment, and remedying its discriminatory effects. All university employees have reporting responsibilities to the Office of Institutional Equity to ensure the university can take appropriate action:

* All university employees, except those exempted by legal privilege of confidentiality or expressly identified as a confidential reporter, have an obligation to report incidents of sexual assault immediately.
* The following employees have an obligation to report all other forms of sexual misconduct as soon as practicable but at most within five workdays of becoming aware of such information: 1. Any human resource professional (HRP); 2. Anyone who supervises faculty, staff, students, or volunteers; 3. Chair/director; and 4. Faculty member.

### Commitment to a diverse and inclusive learning environment

The Ohio State University affirms the importance and value of diversity in the student body. Our programs and curricula reflect our multicultural society and global economy and seek to provide opportunities for students to learn more about persons who are different from them. We are committed to maintaining a community that recognizes and values the inherent worth and dignity of every person; fosters sensitivity, understanding, and mutual respect among each member of our community; and encourages each individual to strive to reach his or her own potential. Discrimination against any individual based upon protected status, which is defined as age, color, disability, gender identity or expression, national origin, race, religion, sex, sexual orientation, or veteran status, is prohibited.

### Your mental health

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student’s ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you or someone you know are suffering from any of the aforementioned conditions, you can learn more about the broad range of confidential mental health services available on campus via the Office of Student Life’s Counseling and Consultation Service (CCS) by visiting [ccs.osu.edu](http://ccs.osu.edu/) or calling 614­-292-­5766. CCS is located on the 4th Floor of the Younkin Success Center and 10th Floor of Lincoln Tower. You can reach an on-call counselor when CCS is closed at 614­-292-­5766 and 24 hour emergency help is also available 24/7 by dialing 988 to reach the Suicide and Crisis Lifeline.

### Accessibility accommodations for students with disabilities

#### Requesting accommodations

The university strives to maintain a healthy and accessible environment to support student learning in and out of the classroom. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion.

If you are ill and need to miss class, including if you are staying home and away from others while experiencing symptoms of viral infection or fever, please let me know immediately. In cases where illness interacts with an underlying medical condition, please consult with Student Life Disability Services to request reasonable accommodations. You can connect with them at slds@osu.edu; 614-292-3307; or slds.osu.edu.

### Religious accommodations

Ohio State has had a longstanding practice of making reasonable academic accommodations for students’ religious beliefs and practices in accordance with applicable law. In 2023, Ohio State updated its practice to align with new state legislation. Under this new provision, students must be in early communication with their instructors regarding any known accommodation requests for religious beliefs and practices, providing notice of specific dates for which they request alternative accommodations within 14 days after the first instructional day of the course. Instructors in turn shall not question the sincerity of a student’s religious or spiritual belief system in reviewing such requests and shall keep requests for accommodations confidential.

With sufficient notice, instructors will provide students with reasonable alternative accommodations with regard to examinations and other academic requirements with respect to students’ sincerely held religious beliefs and practices by allowing up to three absences each semester for the student to attend or participate in religious activities. Examples of religious accommodations can include, but are not limited to, rescheduling an exam, altering the time of a student’s presentation, allowing make-up assignments to substitute for missed class work, or flexibility in due dates or research responsibilities. If concerns arise about a requested accommodation, instructors are to consult their tenure initiating unit head for assistance.

A student’s request for time off shall be provided if the student’s sincerely held religious belief or practice severely affects the student’s ability to take an exam or meet an academic requirement **and** the student has notified their instructor, in writing during the first 14 days after the course begins, of the date of each absence. Although students are required to provide notice within the first 14 days after a course begins, instructors are strongly encouraged to work with the student to provide a reasonable accommodation if a request is made outside the notice period. A student may not be penalized for an absence approved under this policy.

If students have questions or disputes related to academic accommodations, they should contact their course instructor, and then their department or college office. For questions or to report discrimination or harassment based on religion, individuals should contact the Office of Institutional Equity.

Policy: [Religious Holidays, Holy Days and Observances](https://oaa.osu.edu/religious-holidays-holy-days-and-observances)