

Southern Illinois University Edwardsville
STAT 562-Machine Learning and Classification Methods
Syllabus – Fall 2023, 8 Week Course

Instructor: Beidi Qiang

Office Location: SE2356

Instructor email: bqiang@siue.edu

Office Hours: Wednesday 1-3pm online via Zoom or by appointment.

Zoom Link: <https://siue.zoom.us/j/95696159593>

Textbook:

An Introduction to Statistical Learning with Applications in R by James, Witten, Hastie, and Tibshirani, 2nd edition, Springer 2021

The pdf version for this book is available for free on the book website.

<https://www.statlearning.com/>

Modern Data Science with R by Baumer, Kaplan, and Horton, CRC Press 2017

The digital edition is available here for free.

<https://mdsr-book.github.io/mdsr2e/index.html#nd-edition>

Course Description: This course provides an overview of advanced classification methods and common supervised and unsupervised machine learning techniques. Topics include classification methods, tree-bases models, ensembled methods, black-box learning models and clustering. Students will also learn model selection and validation using training and testing and cross-validation. Computing methods will be illustrated using R.

Course Objectives: After the completion of this class, students will be able to:

- Employ common supervised and unsupervised learning techniques.
- Use classification methods on supervised data.
- Use flexible tree-based prediction models.
- Use ensemble prediction models.
- Perform model tuning and selection.

Course Outline and Topics:

- Linear and quadratic discriminant analysis
- Nearest-Neighbors
- Tree-based methods (Decision Trees)
- Support-vector machines
- Ensemble methods (bagging and random forest)
- Deep learning (artificial neural nets and convolutional neural nets)
- Generating training and validation datasets for cross-validation
- Bootstrapping and resampling methods
- Regularization
- Unsupervised learning (clustering)

Blackboard: Supplemental resources (PPT slides, tables, etc) will be posted on Blackboard and can be accessed at (<http://bb.siue.edu>). Your Blackboard username and password is the same as your SIUE e-id and password.

Software: You will learn to use the software package R, which is a free, open-source statistical programming language. Details on how to download R for free onto your home computer and how to get started with R can be found at (<https://goo.gl/Jfr3mS>) or at the R Project Site (<http://cran.r-project.org/>).

Grading Scale: Your final grade in this course will be determined as follows

4 Homework	- 40%
1 Online Final Exam	- 30%
Course Project	- 30%

The straight scale will be used in assigning your final grade, i.e., A: 90 - 100%, B: 80-89%, C: 70-79%, D: 60-69% and F: less than 59%.

Academic integrity: Plagiarism is the use of another person's words or ideas without crediting that person. Plagiarism and cheating will not be tolerated and may lead to failure on an assignment, in the class, or dismissal from the University, per the [SIUE academic dishonesty policy](#). Students are responsible for complying with University policies about academic honesty as stated in the [University's Student Academic Conduct Code](#).

Students Needing Accommodations: Students needing accommodations because of medical diagnosis or major life impairment will need to register with Accessible Campus Community & Equitable Student Support (ACCESS) and complete an intake process before accommodations will be given. Students who believe they have a diagnosis but do not have documentation should contact ACCESS for assistance and/or appropriate referral. The ACCESS office is located in the Student Success Center, Room 1270. You can also reach the office by e-mail at myaccess@siue.edu or by calling 618.650.3726. For more information on policies, procedures, or necessary forms, please visit the ACCESS website at www.siue.edu/access.

Technology Requirements:

Technical requirements for students can be found in this [ITS Knowledge Base article](#).

Technology capabilities: Students in an online course should be able to:

- Use a word processor, such as MS Word, to compose assignments and communicate with others in class.
- Attach files to emails or course areas.
- Navigate websites and course materials.
- Reach out to tech support staff when issues arise and troubleshoot to resolve problems.

Additional guidance for taking online courses can be found on the [Online at SIUE site](#).

Technical Support:

Since this is an online course, you are expected to have reliable Internet access on a regular basis. It is your responsibility to address any computer problems that might occur. Such problems are not an excuse for delays in meeting expectations or for missing course deadlines.

Contact ITS at [618-650-5500](tel:618-650-5500) or at help@siue.edu with any technical concerns. You can also check the functionality of University systems, including Blackboard, at the [ITS System Status page](#), or search the [ITS Knowledge Base](#) for various how-to and troubleshooting guides.

Tips for taking online assessments:

- Set up a wired (Ethernet) Internet connection on your computer.
- Do not use a mobile device, such as a phone or tablet.
- Read the instructions and directions carefully.

Be prepared to complete the assessment in the allotted time.

Course Schedule:

Week 1	Introduction to Statistical Learning
Week 2	Classification Methods
Week 3	Cross Validation and Regularization
Week 4	Tree-based Methods
Week 5	Ensemble methods (bagging and random forest)
Week 6	Supported Vector Machine
Week 7	Artificial Neural Networks and Deep learning
Week 8	Unsupervised learning, Clustering

Subject to change notice:

All material, assignments, and deadlines are subject to change with prior notice. It is your responsibility to stay in touch with your instructor, review the course site regularly, or communicate with other students, to adjust as needed if assignments or due dates change.