

MATH 4385/5385 Statistical Learning

What will students learn from the course?

As part of the [Graduate Certificate in Data Analytics program](#), we will teach classical machine learning and statistical methods with hands-on R coding skills for data analysis of real applications. The **main topics** include bias-variance trade-off, multiple linear regression, classification (logistic regression, linear discriminant analysis, KNN, Naïve Bayes), resampling methods (bootstrap and K-fold CV), model selection, regularization (lasso, ridge, elastic net), principal component regression, non-linear models, tree methods (random forest, boosting), support vector machines, and unsupervised learning methods.

Students will learn how these methods work, how to formulate questions from real problems, select appropriate methods, run analysis and modeling in **R**, interpret and present analysis results in a visualizable, professional, and insightful report.

What is the prerequisite and what software and tools will be taught and used?

The **prerequisite** is **Math 2314** (Elementary Statistical Methods). The course will introduce and use the free open-source statistical software **R**, the integrated development environment **RStudio**, and the **R Markdown** tool. All lecture notes and projects will be written and provided in **R Markdown** documents. No prior experience of **R** is required.

What is the main textbook?

We will use the book with its corresponding **R** package **ISLR2**: James, G., Witten, D., Hastie, T. and Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*, 2021, 2nd edition, Springer, NY. (ISBN 978-1071614174) (downloadable at <https://www.statlearning.com/>)

How the course is taught?

It is a 100% **online** class **offered in each fall semester**. Weekly reading materials will be on **Canvas** including lecture **notes**, **videos**, and source files. Students manage time on their own. Four assignments (written homework + **R** projects), midterm exam, final exam, and final project will be the major submissions. **Online platform** on **Piazza.com** will be the forum for communications among students and the instructor.

Current instructor

Xueyan Sherry Liu, xliu10@uno.edu, Department of Mathematics at UNO.

