

MATH 6375 Advanced Statistical Learning

What will students learn from the course?

The course is a part of the [Graduate Certificate in Data Analytics program](#). We discuss some major statistical learning methods at an advanced level with hands-on R coding skills for more advanced real data applications. Main topics include unsupervised learning, penalized regression, model assessment and selection, model inference and averaging, additive and tree methods, neural networks, high-dimensional problems, and etc. The focus of the class will be on in-depth instruction of the statistical concepts and related analytical skills using **R**. 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 4301/5301, or 4304/5304, or 4311/5311, or 4312/5312. The first course Math 4385/5385 (Statistical Learning) is not listed and you could directly take Math 6375 without taking 4385/5385 in advance. The course will introduce and use the free open-source statistical software **R**, the IDE **RStudio**, and the **R Markdown** package. All lecture notes and projects will be written and provided in **R Markdown** documents. No prior experience of **R** is required, as well.

What is the main textbook?

We will use part of the following book with its corresponding **R** package **ElemStatLearn**: Hastie, Tibshirani, Friedman. *The Elements of Statistical Learning*, 2009, 2th edition, Springer, New York. (ISBN 978-0-387-84858-7) (downloadable at <https://hastie.su.domains/ElemStatLearn/>)

How the course is taught?

The course is 100% online and offered in spring semesters. How the course is taught is similar to Math 4385/5385. Check the flyer of Math 4385/5385 for detail.

What is the relation between Math 4385/5385 and Math 6375?

The Math 4385/5385 is the introductory level course of statistical learning and Math 6375 is more in-depth in theory and more advanced. Ideally, students are suggested to take Math 4385/5385 first and then Math 6375. However, students with background of senior or graduate level of statistics and mathematics (such as, math stat, linear algebra, etc.) could take Math 6375 without Math 4385/5385.

Current instructor:

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

