

MATH 567: Mathematical Techniques in Data Science

Time & Place: MW 5:00PM – 6:15PM, Room: Ewing Hall (EWG) 205.

Instructor: Dominique Guillot, dguillot@udel.edu, Office: Ewing Hall 534.

Office hours details: Monday 11:15–12:15 + Thursday 2:30–3:30 + by appointment.

Venue: Ewing Hall 534.

Course description: The course provides an introduction to the fundamental techniques used in data mining. The main objective of the course is to develop a good practical knowledge and a mathematical understanding of the common tools that are used to analyse modern datasets. The course also provides hands-on experience in data analysis through practical homework and class projects.

Goals of the course:

- Become familiar with the basic methods used to analyse modern datasets.
- Understand some of the mathematical properties of standard techniques in data mining.
- Understand how to select a good model for data.
- Be able to analyse datasets using a modern programming language such as Python or R.

List of topics:

Linear methods for regression (subset selection, ridge, lasso), Logistic regression. Linear discriminant analysis, Principal component analysis, Density estimation, Decision trees, Random forests, Neural networks, Support Vector Machines, Cluster analysis (K-means, spectral clustering), Network models, Cross-validation, Bootstrap.

Textbook:

An Introduction to Statistical Learning by James, Witten, Hastie, and Tibshirani.

Assessment plan:

- Midterm (25%), Wednesday March 22 (in class);
- Final (25%), Date/time TBA;
- Class project (20%);
- Homework + Labs (30%).

Grading system: $A > 94\%$, $A- \geq 90\%$, $B- \geq 80\%$, $C- \geq 70\%$, $D- \geq 60\%$, $F < 60\%$. Subject to change (to your advantage only).

Prerequisites

Probability theory and basic statistics (e.g. MATH 350), Multivariable calculus (e.g. MATH 222), Linear Algebra (e.g. MATH 349), Optimization background (e.g. MATH 529) desirable but not necessary, basic computing skills.

Classroom Conduct: As a common courtesy to the instructor and your fellow classmates, please refrain from using your cell phone and computer in the classroom (unless instructed otherwise).

Academic Integrity Statement: All University of Delaware policies regarding ethics and honorable behavior apply to this course. Cheating receives a failing grade. For more details, please refer to the *Student Guide to University Policies* available at: <http://www.udel.edu/stuguide/15-16/code.html>.

Late homework policy: Late homework will not be accepted unless prior arrangements are made with the instructor.

Faculty Statement on Disclosures of Instances of Sexual Misconduct Please see the following link for information regarding the University's policy about sexual misconduct.