

- **Instructor:** Dr. Marc G. Genton  
E-mail: genton@stat.tamu.edu  
Phone: 845-3152  
Office: 405D Blocker  
Office hours: 3:45-5:00PM Tuesday and Thursday
- **Preq:** STAT 610/611
- **Course Schedule:** T R: 9:35-10:50AM in Blocker 411
- **Web page for STAT 689:**  
<http://www.stat.tamu.edu/~genton/STAT689/STAT689.html>
- **Course Objectives:** By the end of the course, the student will be able to:
  - (1) Understand and describe various methods of data mining.
  - (2) Apply various methods of data mining to real data sets.
  - (3) Use the software *S-PLUS* (or *R*, or *SAS*) to perform data mining on real data sets.
- **Course Outline:**
  - **Supervised learning:** linear methods for regression/classification; additive models and trees (GAM, CART, PRIM, MARS, MART); bagging and boosting methods; neural networks; support vector machines.
  - **Unsupervised learning:** association rules; market basket analysis; cluster analysis; independent component analysis.
- **Textbook (required):**
  - Hastie, T., Tibshirani, R., Friedman, J. (2001), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- **Course work:** Your grade for the course will be based on (relative weights given in %):
  - **Homework assignments (50%):** Bi-weekly homework assignments will be given throughout the semester. Assignments will be collected at the START of class on the date due. Late assignments will not be accepted.

- **Project (50%):** A project, done individually, will be due near the end of the semester, and presented in class (20 minutes). More details will be given as the semester progresses. Late projects will not be accepted.

The following letter grades are guaranteed:

A	B	C	D	F
100-90%	90-80%	80-70%	70-60%	60-0%

- **Notes:**

- Check the website frequently for updates!
- I would like to know about any particular academic difficulties or personal problems that are affecting a student's performance.

January 7, 2005. Marc G. Genton