

Analytical Informatics

*Better analytical chemistry
through computers?*

Part I. A Statistical Foundation for Informatics

- ◆ What is analytical informatics?
- ◆ Populations, features and classes
- ◆ Random variables and linear transformations
- ◆ Practical statistical tools: the R environment
- ◆ Introduction to statistical decision theory
- ◆ Simple prediction: least squares
- ◆ Simple prediction: nearest neighbors

Part II. Applied Machine Learning and Pattern Recognition

- ◆ Linear methods for classification
- ◆ Feature extraction and dimensions
- ◆ Model evaluation
- ◆ Neural Networks
- ◆ Support Vector Machines
- ◆ Tree-based methods
- ◆ Ensemble methods

Part III. Technology for Chemical Data Mining and Predictive Modeling

- ◆ Chemical data representation
- ◆ Databases
- ◆ Programming for database mining
- ◆ Integrating applications: Perl, R and MySQL
- ◆ Building and testing predictive models
- ◆ Putting it all together

Technology Interludes

- ◆ Computer languages
- ◆ High performance computing and networks
- ◆ Bioinformatics technology
- ◆ Data Access

Important Dates

- ◆ Jan 17: Last day for late registration
- ◆ Jan 27: Last day to cancel with no record
- ◆ Feb 10: Last day to cancel course without a grade
- ◆ Feb 10-17: Exam 1
- ◆ March 17-22: Spring Break
- ◆ March 24: Last day to cancel with Pass/Fail record
- ◆ March 24-31: Exam 2
- ◆ April 21-30: Exam 3
- ◆ May 3: Semester ends
- ◆ May 13: Grades due to registrar