ISyE/CSE 6740 Computational Data Analysis/Machine Learning Syllabus

Theoretical/computational foundations of analyzing large/complex modern datasets, including the fundamental concepts of machine learning and data mining needed for both resesarch and practice. Crosslisted with CSE 6740.

Course Objective: We will discuss a number of most commonly used machine learning models and algorithms in sufficient amount of details on their mechanisms.

Classes: 15 weeks, 3 hours per week (1:30pm-2:45pm T, Th, Instr Center 211)

Instructor: Guanghui (George) Lan, Office: Groseclose 445

Prerequisites: 1. Probabilities, 2.Statistics, 3. Linear algebra, 4. Algorithms and Programming, 5. Convex optimization.
(1-4) are real pre-requisites. Basics of convex optimization will be covered in the course.

Grading policy: homework (20%), two midterms (20% each), Final exam (40%).

References (not required):

1. T. Hastie, R. Tibshirani, J. Friedman, the Elements of Statistical Learning.

2. G. Lan, Lectures on Optimization Methods for Machine Learning.

Tentative exam dates: 9/24 (Midterm 1), 11/05 (Midterm 2), 12/05-12/12 (Final)

Tentative schedule:

- Unit#1: Unsupervised learning (data exploration)

    - Clustering vectorial data
    - Clustering networks
    - Dimensionality reduction
    - Dimensionality for manifold
    - Density estimation

- Unit#2: Supervised learning (predictive models)

    - Classifications

- Combined classifiers
- Regressions

- Unit#3: Complex models (dealing with nonlinearity, combine models etc.)

  - Kernel methods
  - Hidden Markov models
  - Graphical models
  - Topic modeling
  - Social network analysis
  - Collaborative filtering