Machine Learning for Math-to-Industry Bootcamp

Summer 2017

Instructor: Gilad Lerman
lerman@umn.edu
612-624-5541
534 Vincent Hall

Time: June 29-30, 9am-noon
Location: Lind Hall, Room 305

Course description and objectives

The short tutorial will introduce students to machine learning methods that are common in industry and that their basic ideas can be understood in a relatively short time.

Brief Syllabus

- K-nearest neighbors as a simple solution for some problems in machine learning
- Support vector machines: maximal margin classifier, support vector classifiers, non-linear decision boundaries and the kernel trick
- Tree-based methods: regression and classification trees, bagging, random forests, boosting
- Principal component analysis: quick review of singular value decomposition, statistical interpretation and applications.
- Two examples of nonlinear methods: SNE/tSNE and SOM
- Clustering techniques: k-means, mixture of Gaussians, hierarchical, DBSCAN, nearest neighbors, spectral

Main Reference (with additional material provided in class)

An Introduction to Statistical Learning *with Applications in R* by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (June 2013). Book Homepage pdf (9.4Mb, 6th corrected printing)