MSSC 6250 — SPRING 2022

# Statistical Machine Learning

Gareth James
Daniela Witten
Trevor Hastie
Robert Tibshirani

An Introduction to Statistical Learning

with Applications in R

Tentative topics :

- **Multivariate Data and Exploratory Analysis**
- **Multivariate Normal Distribution**
- **Supervised Learning**:
  - Classification and Regression
    - Logistic Regression and Linear Discriminant Analysis
    - Ridge Regression and Lasso
  - Tree Based Methods
    - Classification tree and Regression tree
    - Bagging, Random Forests, Boosting
  - Support Vector Machine
- **Unsupervised Learning** (Dim. Reduction and Clustering)
  - Principal Component Analysis (PCA)
    - Human genetic clustering
    - Google PageRank
  - Independent Component Analysis (ICA)
    - Blind Source Separation
  - K-Means and Hierarchical Clustering
- **Moving Beyond Linearity** (Smoothing Splines)
- **Natural Language Processing** (NLP)
  - Bag of Words; TF-IDF; Word2Vec
- **Neural Networks**
  - Introduction to Deep Learning
    - TensorFlow and Keras, *or*
    - PyTorch and fast.ai
  - Convolutional Neural Networks (CNN)
  - Recurrent Neural Networks (RNN)
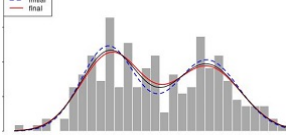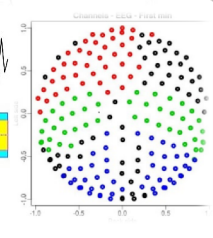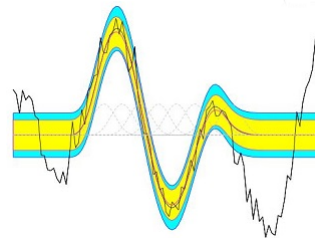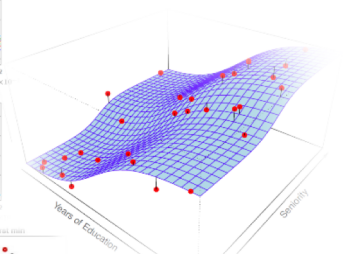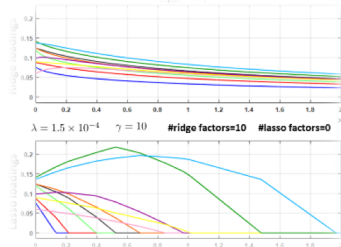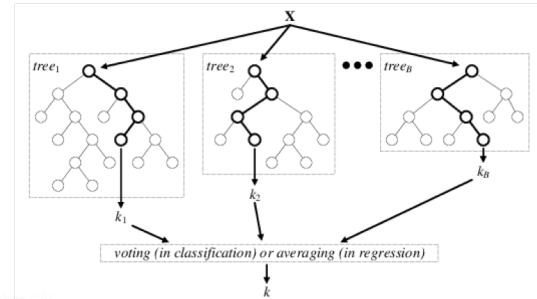
Prerequisites:

- A course in Statistical Methods
- A course in Linear Algebra

For more information, email the instructor:

- Dr. Mehdi Maadooliat (mehdi.maadooliat@marquette.edu)