Instructor: Tuo Zhao

Time and Location: TTh 3:00-4:15pm, Clough Commons 152

Contact: tourzhao@gatech.edu

Office Hours: Th 1:45-2:45pm, Groseclose 344

TAs: Minshuo Chen, Shaojun Ma, Yujia Xie, Yu Cao, Zhehui Chen and Haoming Jiang

References

• The Elements of Statistical Learning, 2nd Edition, Hastie, Tibshirani and Friedman.
• Foundations of Machine Learning, Mohri, Rostamizadeh and Talwalker
• Andrew Ng. CS229 Lecture Notes: http://cs229.stanford.edu/materials.html

Honor code

For any questions involving Academic Honor Code issues, please consult me, the class teaching assistant or http://www.honor.gatech.edu.

Background Knowledge Test (6%)

Time: Day One. The exam will be 1 hour long, and closed to books and notes, and no electronic device (e.g. cell phone, laptop) is allowed.

A make-up exam is scheduled on Sep. 19. The grades will be curved according to the background knowledge test.

Exam-1 (26%)

Time: Oct 14. The exam will be 1.25 hour long, and closed to books and notes, and no electronic device (e.g. cell phone, laptop) is allowed.

Exam-2 (26%)

Time: Nov. 30. The exam will be 1.25 hour long, and closed to books and notes, and no electronic device (e.g. cell phone, laptop) is allowed.
Homework (42%)

7 homework assignments. Each homework assignment consists of either a few analytical problems or simple coding problems. You are required to use LaTeX for your analytical HW, and MATLAB for your coding HW.

Covered Topics (Theory:Methodology:Algorithm=2:3:2)

- General Introduction
- Supervised Learning, Discriminative Algorithms:
  Supervised Learning Concept, Linear Regression, Maximum Likelihood, Normal Equation
  Gradient Descent, Stochastic Gradient, SVRG.
  Linear Classification, Logistic Regression, Newton Method,
- Generative Algorithms:
  Multivariate Normal, Linear Discriminant Analysis
  Naive Bayes, Laplacian Smoothing
  Multiclass Classification, K-NN
  Multi-class Fisher Discriminant Analysis, Multinomial Regression
  Support Vector Machines and Kernel Methods:
  Intuition, Geometric Margins, Optimal Margin Classifier
  Kernel, SMO algorithm, Coordinate Gradient Descent.
  Kernel PCA, Kernel Logistic Regression, Kernel Ridge Regression, Multiclass SVM
- Unsupervised Learning:
  PCA, Mixture Models, Bayesian Graphical Models
  Power Method, Oja’s algorithm, EM Algorithm, Variational Inference
  Matrix Factorization/Completion
- Regularization and Model Selection:
  Cross Validation, Hill Climbing, Bayesian Optimization
  Bayesian Regression, Bayesian Logistic Regression
  Forward and Backward Regression, Lasso, elastic-net.
  Proximal Gradient, Prox-SVRG.
  Coordinate Proximal Gradient, Pathwise Coordinate Descent
- Decision Tree and Random Forest:
  Entropy, Building Tree
  Bagging features, Bagging Samples, Random Forest
  Adaboost, Gradient Tree Boosting
• Neural Network:
  Concept; Deep Neural Network; Backpropagation
  Convolutional Neural Network;