Stat 231 / CS 276A
Pattern Recognition and Machine Learning

Fall 2017, MW 2:00-3:15 PM, Physics and Astronomy Building 1434A
www.stat.ucla.edu/~sczhu/Courses/UCLA/Stat_231/Stat_231.html

Course Description

This course introduces fundamental concepts, theories, and algorithms for pattern recognition and machine learning, which are used in computer vision, speech recognition, data mining, statistics, information retrieval, and bioinformatics. Topics include: Bayesian decision theory, parametric and non-parametric learning, data clustering, component analysis, boosting techniques, support vector machine, and deep learning with neural networks.

Prerequisites

You don't have to take exactly these courses as long as you know the materials.

- Math 33A Linear Algebra and Its Applications, Matrix Analysis.
- Stat 100B Intro to Mathematical Statistics.
- CS 180 Intro to Algorithms and Complexity.
- Programming skills in Matlab or Python.

Textbook

Textbook is not mandatory if you can understand the lecture notes and handouts.

- R. Duda, et al., *Pattern Classification*, John Wiley & Sons, 2001. [Good for CS students]
- T. Hastie, et al., *The Elements of Statistical Learning*, Spinger, 2009. [Good for Stat students]
- C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. [with advanced materials]

Instructors

- Prof. Song-Chun Zhu, sczhu@stat.ucla.edu, Office: Boelter Hall 9404
  Office Hours: Tuesday 1-3pm
- Reader: Ruiqi Gao for stat 231, ruiqigao@ucla.edu, Erik Nijkamp for CS 276A, enijkamp@ucla.edu,
  Office Hours: BH 2444, Session 1: Tuesday 6-7pm and Session 2: 7-8pm for project overview and assistant.
  Individual questions will be answered by the readers through CCLE online.

Grading Plan

4 units, letter grades.

Five projects:	
• **0. Exercise on Classification using Deep Learning and CNN** **Experience the training and testing of problems in machine learning**	10%
• **1. Face modeling by AAM Model and FLD** **How many bits do you need to represent a face? Or telling apart male from female faces?**	15%
• **2. Face detection by AdaBoost and RealBoost** **How many features do you need to detect a face in a crowd?**	15%
• **3. Object detection by Fast-RCNN** **How to detect the class of an object and its location?**	10%
• **4. Face social attributes and Political Election Analysis by SVM** **How do we measure the social dimensions of faces in political elections and social network?**	10%

Final Exam: (date and room to be announced)	40%

Grading Policy

- You are encouraged to work and discuss in a group, but each person must finish his/her own project. Submit your report of the project, and your code through the CCLE website.
- You have a total of 5 late days (not including weekends) for the entire class (5 projects) to cover your various reasons, but after using the four late days, no credit will be given for late homework/project.

Schedule

Tentative Schedule for 2017 (Once the enrollment is fixed, course materials will be posted on the CCLE site.)

Lecture	Date	Topics	Handouts
1	10-02	**Introduction to Pattern Recognition** [Problems, applications, examples]	
2	10-04	**Bayesian Decision Theory I** [Bayes rule, discriminant functions]	
3	10-09	**Bayesian Decision Theory II** [loss functions and Bayesian error analysis]	
4	10-11	**Component Analysis and Dimension Reduction I**: [PCA, face modeling by Active Appearance Model (AAM)]	
5	10-16	**Component Analysis and Dimension Reduction II** [Fisher Linear Discriminant] [Multi-dimensional scaling (MDS)]	
6	10-18	**Component Analysis and Dimension Reduction III** [Local Linear Embedding (LLE), Intrinsic dimension]	
7	10-23	**Boosting Techniques I** [perceptron, backpropagation and Adaboost]	
8	10-25	**Boosting Techniques II** [RealBoost and Example on face detection]	
9	10-30	**Boosting Techniques III** [analysis, logit boost, cascade and decision policy]	
10	11-01	**Boosting Techniques III** [analysis, logit boost, cascade and decision policy]	
11	11-06	**Non-metric method I** [Decision tree and random forrest]	
12	11-08	**Non-metric method II** [Syntactic pattern recognition and example on human parsing]	
13	11-13	**Support vector machine I** [Kernel-induced feature space]	
14	11-15	**Support vector machine II** [Support vector classifier]	
15	11-20	**Support vector machine III** [Loss functions, Latent SVM]	
16	11-22	**Parametric Learning** [Maximum Likelihood Estimation] [Sufficient Statistics and Maximum entropy]	
17	11-27	**Non-parametric Learning I** [Parzen window and K-nn classifer]	
18	11-29	**Non-parametric Learning II** [K-nn classifer and Error analysis]	
19	12-04	**Deep Learning I**	
20	12-06	**Deep Learning II**	