Stat 231--- CS 276A

Pattern Recognition and Machine Learning

MW 9:30-10:50 Am, Fall 2011, Math Science 5128

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, kernel methods and support vector machine, and fast nearest neighbor indexing and hashing.

Prerequisites

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

Textbook

- R. Duda, P. Hart, D. Stork, "*Pattern Classification*", second edition, 2000. [Good for CS students] [link to book page]
- T. Hastie, R. Tibshurani, and J.H. Friedman, "*The Elements of Statistical Learning: Data Mining, Inference, and Prediction*", Spinger Series in Statistics, 2001. [Good for Statistics students]

Instructors

- Prof. Song-Chun Zhu, sczhu@stat.ucla.edu, 310-206-8693, office BH 9404.
  Office Hours: Monday 3-4:30pm
- Reader: Brandon Rothrock, rothrock@cs.ucla.edu, office BH 9410,
  Office hours: Thursday 4-6:00pm

Grading Plan: 4 units, letter grades

Two Homework assignments	20%
Three projects: • **1. Face modeling by AAM Model** **how many bits do you need to represent a face?** • **2. Face detection by AdaBoost** **how many features do you need to detect a face? detailed description is here.** • **3, Object classification by SVM** **classifying objects in video surveillance**	15% 15% 15%
No Middle-Term Exam	0%
Final Exam	35%

Grading policy

- **Homework policy:**
  Homework must be finished independently. Do not discuss with classmates.
- **Project policy**:
  You are encouraged to work and discuss in a group, but each person must finish his/her own project. Hand in (i) a brief description of the experiment in hard copy, (ii) results and plots in hard copy, (iii) your code in e-copy to the reader.

- **Late policy:**
  You have a total of three late days for the class, but after using the three late days, no credit will be given for late homework/project.

Tentative Schedule for 2011

Lecture	Date	Topics	Reading Materials	Handouts
1	09-26	**Introduction to Pattern Recognition** [Problems, applications, examples, and project introduction]	Ch 1	syllabus.pdf Lect1.pdf
2	09-28	**Bayesian Decision Theory I** [Bayes rule, discriminant functions]	Ch 2.1-2.6	Lect 2.pdf
3	10-03	**Bayesian Decision Theory II** [loss functions and Bayesian error analysis]	Ch 2.1-2.6	Lect 3.pdf
4	10-05	**Component Analysis and Dimension Reduction I**: [principal component analysis (PCA)], face modeling] [Explanation of Project 1: code and data format]	Ch 3.8.1, Ch 10.13.1 Project 1	HW1 Lect4-5.pdf
5	10-10	**Component Analysis and Dimension Reduction II:** [Fisher Linear Discriminant] [Multi-dimensional scaling (MDS)]	Ch 3.8.2, Ch10.14	FisherFace.pdf Lect5-6.pdf
6	10-12	**Component Analysis and Dimension Reduction III:** [Local Linear Embedding (LLE), Intrinsic dimension]	paper	LLE paper
7	10-17	**Boosting Techniques I**: [Adaboost]	Ch 9.5	Lect7-9.pdf
8	10-19	**Boosting Techniques II**: [RealBoost and Example on face detection] [Explanation of project II]	Tutorial Handout 1 Handout 2	
9	10-24	**Boosting Techniques III**: [Probabilistic analysis, Logit boost]		
10	10-26	**Non-metric method I**: [tree structured Classification: principle and example]	Ch 8.1-8.3	Lect10.pdf
11	10-31	**Non-metric method II**: Syntactic pattern recognition	Ch 8.5-8.8	Lect11.pdf
12	11-02	**Support vector machine I**: Kernel-induced feature space	Tutorial paper	Lect12-15.pdf (xerox handout)
13	11-07	**Support vector machine II:** [Support vector classifier] [Explanation of project III]	Ch 5.11	
14	11-09	**Support vector machine III:** [Loss functions, Latent SVM]		
15	11-14	**Parametric Learning** [Maximum Likelihood Estimation (MLE)] [Sufficient Statistics and Maximum entropy]	Ch 3.1-3.6	Lect16.pdf

16	11-16	**Non-parametric Learning I** [Parzen window]	Ch 4.1-4.5	Lect17.pdf
17	11-21	**Non-parametric Learning II**: [K-nn classifer and Error analysis]	Ch 4.6 handout	Lect18.pdf
18	11-23	**Non-parametric Learning III:** [K-nn fast approximate computing: KD-tree and Hashing]	paper1 paper2	Lect 19.pdf
19	11-28	**Data Clustering I**: [K-mean clustering, EM]	Ch 10.1-10.4	Lect 20.pdf
20	11-30	**Data Clustering II**: [EM, mean-shift]	Handout	