Syllabus for CSCI 567 Machine Learning

The content of this syllabus is tentative and subject to change.

Introduction: The chief objective is to teach statistical methods in pattern classification and machine learning. Topics include but not limited to various parametric and nonparametric methods for supervised and unsupervised learning problems. Particular focuses on the theoretical understanding of these methods, as well as their computational implications.

Prerequisite: Undergraduate level training or coursework in linear algebra, multivariate calculus, basic probability and statistics; an undergraduate level course in Artificial Intelligence may be helpful but is not required.

Special note: during the first meeting of the class, a 25-minute special quiz will be administered for students to examine their readiness. Students who have expectation of certain grades or above (for instance, in order to improve their GPAs) should exercise their cautions in taking this course, if the quiz appears challenging and leads to a less ideal initial assessment. Exact details will be given at the time of the quiz. Please do come to the first meeting if you intend to take the course.

Format: classroom lectures, homework, in-class quizzes (2 or 3) and mini-projects

Grading: Based on quizzes and assignments. Relative weights will be decided by the instructor and may change from semester to semester. It is anticipated that initially assignments and projects will count for 60% of the total; quizzes for the rest.

Instructors: Initially, the course will be taught by Prof. Sha or Prof. Nevatia. However, many faculty in computer science are experts in the topic and teaching may rotate among these professors. The contents and the focus will vary.

Schedule: It is proposed to offer this course every fall semester.

Textbooks:
Required:
- Pattern recognition and machine learning (by Christopher Bishop),
- chapters of textbooks in preparation (permission of the authors will be sought).
Optional:
- Elements of Statistical Learning (Trevor Hastie, Robert Tibshirani, and Jerome Friedman)
- Pattern Classification (Richard O. Duda, Peter E. Hart, and David G. Stork)
- Introduction to Artificial Intelligence (by Stuar Russell and Peter Norvig)
- Machine Learning (by Tom Mitchell)

Schedule
Introduction (~ 1.5 weeks)
- History and broad overview of machine learning
- Review of mathematical tools: linear algebra, multivariate calculus, probability and statistics, optimization
Brief introduction of programming environment: Matlab and R.

Supervised Learning (~5 weeks)
   General notions (Bayes optimality, curse of dimensionality, overfitting and model selection, bias vs. variance tradeoff, generative vs. discriminative for parameter estimation, feature selection, and etc)
   Linear methods (linear, logistic regression and generalized linear models, naive Bayes, linear discriminant analysis, support vector machines, and etc)
   Nonlinear methods (kernel methods, nearest neighbor, decision trees, neural networks, and etc)
   Ensemble learning (bagging, boosting, and etc)

Unsupervised learning (~4.5 weeks)
   Clustering and density estimations (K-means/vector quantization, mixture models, etc)
   Dimensionality reduction (linear and nonlinear methods)

Advanced topics (~3 weeks)
   Basic learning theory
   Various topics (e.g., Bayesian approaches, semi-supervised learning, online learning, sparsity and etc)

Statement for Students with Disabilities
Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

Statement on Academic Integrity
USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one’s own academic work from misuse by others as well as to avoid using another’s work as one’s own. All students are expected to understand and abide by these principles. Scampus, the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://www.usc.edu/dept/publications/SCAMPUS/gov/. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: http://www.usc.edu/student-affairs/SJACS/.