Introduction The chief objective of this course is to teach methods in pattern classification and machine learning. Key components include statistical learning approaches, including 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.

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

Special note I: After the first meeting of the class on 8/26, a special take-home quiz will be released for students to examine their readiness. The quiz will *not* count toward to the final grade. It will be open-book. It will be due back on Aug 27, 2013 10:30am PDT, to Graduate Advisors' offices at SAL 308.

The quiz will be graded and assessed by the instructor and the TAs. Students who do not meet the passing threshold are *strongly urged to reconsider their decisions* or not permitted to take the course and might need to withdraw from the class. Furthermore, 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.

Please do come to the first meeting if you intend to take the course, whether you were already registered or are still on the waiting list.

Please be advised, to avail yourself a 2-hour time window, between 8/26/2013 7:00pm PDT to 8/27/2013 10:30am PDT, to finish the quiz. The 2-hour is our estimation — you might need less or more.

Programming requirement Students are required to use Matlab for programming exercises/components. Homework assignments are required to be typeset with LaTeX(various TeXeditors and compiling environment on Windows, Mac OS X and Unix/Linux are available, including WYSIWYG ones).

Format classroom lectures, homework, in-class two quizzes. Homework assignments include programming components for algorithmic implementation and mini-projects.

Special note II: students are strongly encouraged to attend a weekly one-hour recitation session (pending scheduling and approval) for in-depth discussion of course material.

Preparation If you would like to prepare or refresh your skills in relevant maths, the followings would be good starting points

- For calculus, please check Prof. Strang's free online textbook

  http://ocw.mit.edu/resources/res-18-001-calculus-online-textbook-spring-2005/textbook

- For linear algebra, please check (again) Prof. Strang's OopenCourseWare site

  http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/index.htm

- Probability and statistics, please check

  http://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statisti

Grading 6 homework assignments (total 60%), two quizzes (total 35%) and class participation (5%).

Policy on homework assignments

- Extension and late turn-in: one two-day extension or two one-day extension for the whole semester; other late turn-in will be penalized with half of the credit.
- Working in group: permitted but each member needs to write up solutions separately.

Teaching Staff instructor (Fei Sha) and TAs: Dong Guo (dongguo@usc.edu) and Zhiyun Lu (zhiyunlu@usc.edu). Office hours will be announced soon.

Required textbooks Kevin Murphy's *Machine Learning: A Probabilistic Perspective*. Additionally, there will be required readings of notes,

Tentative Schedule Please see the last page of this document.

Optional textbooks for references and supplementary reading

- *Elements of Statistical Learning* by Hastie, Tibshirani and Friedman
  http://wwwstat.stanford.edu/~tibs/ElemStatLearn/
- *A course in machine learning* by Hal Daumé III http://ciml.info
- *Bayesian reasoning and machine learning* by David Barber
  http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.HomePage
- *Pattern Recognition and Machine Learning* by C Bishop (available from online and campus bookstores)
- Andrew Moore's Tutorial http://www.autonlab.org/tutorials/
- Andrew Ng's free online course http://ml-class.org/ (started 8/20/2012 for this semester) and lecture material http://cs229.stanford.edu/
- Ben Taskar's Lectures https://alliance.seas.upenn.edu/~cis520/wiki/
- Erik Sudderth's Course and Collection of Resources
  http://www.cs.brown.edu/courses/cs195-5/resources.html
- *Pattern Classification* by Duda, Hart and Stork
- *All of Statistics* by L. Wasserman

References for frequently used maths

- *The Matrix cookbook*
  http://orion.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
- Chris Burges's note on applied maths for machine learning
  http://research.microsoft.com/en-us/um/people/cburges/tech_reports/tr-2004-56.
- The Wisconsin collection http://pages.cs.wisc.edu/~andrzeje/lmml.html
- Khan Academy: http://www.khanacademy.org/

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 ones own academic work from misuse by others as well as to avoid using anothers work as ones 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/.

Date	Topics to be covered	Notes
8/26	Overview of ML; nearest neighbor classification	Special Quiz
8/28	Decision trees; maximum likelihood estimation	
9/2	No class	Labor Day
9/4	Gaussian random variables; linear regressions	HW1 out
9/9	Logistic regression; numerical optimization	
9/11	Generative vs. discriminative; naive Bayes, Gaussian discriminant	
9/16	Overfitting, bias/variance tradeoff, regularization	
9/18	Bayesian approaches: Bayesian linear regression	HW2 out
9/23	Gaussian process; neural networks	
9/25	Kernel methods; SVM; convex duality	
9/30	Generalization theory	
10/2	Boosting	HW3 out
10/7	Quiz 1	
10/9	Clustering, mixture models	
10/14	Mixture models; EM algorithm	
10/16	PCA	HW4 out
10/21	Factor analysis; probabilistic PCA	
10/23	ICA; manifold learning	
10/28	NMF; PLSA	
10/30	Possibly TA-led lectures (Prof. Sha traveling)	HW5 out
11/4	Bayesian network	
11/6	Hidden Markov networks	
11/11	Topic models	
11/13	Variational and sampling methods	HW6 out
11/18	Markov random fields	
11/20	Markov random fields	
11/25	Selected topics: sparse learning, online algorithms, large-scale learning	
11/27	No class	Happy Thanksgiving
12/2	Course review/summary	
12/4	Quiz 2 (cumulative but excluding 11/25 lecture)	