CSCI 467 Fall 2018 Syllabus

Course meetings: You are responsible for everything covered in lectures and your enrolled discussion, including administrative announcements.

Strongly Recommended Textbooks: Tom E. Mitchell, *Machine Learning*
Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, *Introduction to Data Mining, Second Edition*
Christopher Bishop. *Machine Learning and Pattern Recognition*
Also Recommended Textbooks: A "Recommended Textbooks" list is available at the end of this syllabus. These cover both mathematical preliminaries and machine learning topics.
Gradebook: blackboard.usc.edu
Forums: We will use Piazza for most electronic communication.

Blackboard will be used for posting of grades. Messages that do not need a particular instructor's direct attention should be posted to Piazza with the appropriate privacy setting. If you email your instructor, you *must* include the substring "CSCI 467" to begin a *meaningful* subject line and have tried to resolve the issue appropriately otherwise (e.g., questions about course material should be posted to Piazza first, using email only after an appropriate amount of time has passed without a response.) Such emails must be sent from your USC email account. Any emails not conforming to this might not be seen by your instructor and will not be considered to have been sent for administrative purposes.

Course Description
Statistical methods for building intelligent and adaptive systems that improve performance from experiences; Focus on theoretical understanding of these methods and their computational implications.
Prerequisites: CSCI 270 AND Math 225 AND (EE 364 OR Math 407 OR BUAD 310g)

Grade Calculations:

Artifact	Percent
Problem Sets *	20%
Programming Assignments	30%
Midterm Exam (date)	20%
Final Exam (date)	30%

Exams

You will be provided with paper on which to take the exam. Exams will be individual effort and closed-book. You will be allowed one normal-sized paper of handwritten notes, front and back. Any requirements for you to bring to the exam will be announced well in advance.

Students requiring alternate exam arrangements must make such requests within the first two weeks of the term, or as soon as possible after knowing of the conflict or requirement.

The midterm exam may be held in a different room than lecture (TBD), but will be held during lecture time. The final exam room is also TBD.

Academic Conduct

It is expected that any work you submit reflect your own understanding of the material at the time of submission. Since we are not collecting problem sets, that should make it *very easy* for you to abide by this on those. For the programming assignments, please follow the "Kenny Loggins" rule:

The "Kenny Loggins" Rule:

You may discuss high-level ideas, and give hints to other students regarding how to solve the programming assignments. Any time you seek help on, or discuss with someone else, a take-home quiz question that you have yet to solve, do not keep any written record of the discussion. Afterwards, take a 30-minute break and do something unrelated to the course (watching a 30-minute episode of your favorite cartoon show, for example). You may now return to your assignment.

As an elective, CSCI 467 is offered with the expectation that students are taking this because they want to learn the material. As this is the first semester the course is being offered, it is possible that more collaboration may be desired by students than is initially permitted by the rule above. As such, we offer the following amendment to the usual rules for students who don't put the assignments off to the last minute: *if you feel the rules for academic honesty are interfering with your ability to complete an assignment for CSCI 467 because there is conduct you believe should be allowed that is not consistent with the rules as written, send your instructor an email outlining what you want to do and why. It is very likely that the code will be amended to explicitly permit what you wish to do. Please give your instructor 24 hours to respond to the request.*

Academic dishonesty carries significant penalties: the suggested punishment for an infraction is an F in the class. Because of the effort to put this class together, as well as the effort expended to ensure no one needs to cheat, I will take offense to any such behavior.

CSCI 467 Projected Course Schedule: A Weekly Breakdown

Please note that this is a *projected* schedule and is subject to change. Any changes will be announced via the course Piazza page or in lecture.

Textbooks referenced:
[GT] refers to *Algorithm Design* by Michael Goodrich and Roberto Tamassia; this was likely a book used in your algorithms classes (e.g., CSCI 270).
Other references are either the top-listed books ([Mitchell] and [TSKK] and [Bishop]) or are listed in recommended books below.

Date	Topic	Related Reading	Notes
Week 1 Aug 20	Course Overview ML overview Nearest Neighbor classification Core ML concepts; typical steps to developing a ML system	[TSKK] 4.3 [TSKK] 3.4 - 3.8	
Week 2 Aug 27	Concept Learning Decision Trees	[Mitchell] Chapter 2 [TSKK] 4.2 [Mitchell] Chapter 3 [TSKK] 3.3	
No lecture Sep 3 or discussion Sep 3 or 4			
Week 3 Sep 5	Boosting ensemble methods random forests	[TSKK] 4.10 [MLaPP] 16.4.1-16.4.5, 16.4.8, 16.4.9 [ESL] 16.3	
Week 4 Sep 10	Fundamentals: Linear Algebra	[LADW] Ch. 1-6 [EoLA] All	
Week 5 Sep 17	Linear Regression Regression with nonlinear basis Regularized regression Maximum Likelihood Model	[MLaPP] 1.4.5, 7.1-7.3, 7.5.1, 7.5.2, 7.5.4, 7.6 [MLaPP] 1.4.7, 1.4.8 [ESL] 7.1, 7.2, 7.3, 7.10 [Mitchell] Ch 6.4 & 6.5	
Week 6 Sep 24	Generative Models, Naive Bayes Graphical Models and Bayesian Learning	[Mitchell] 6.1-6.3, 6.7-6.10 [TSKK] 4.4 [MLaPP] 3.5 [ESL] 6.6.3 [Bishop 8.1-8.2] [Mitchell] 6.11	

Week 7 Oct 1	Hidden Markov Models Linear discriminant analysis Perceptron	[MLaPP] 17.1-17.4, 17.5.1-17.5.2 [MLaPP] 4.2.1 - 4.2.5, 8.5.1-8.5.4	
Week 8 Oct 8	Review / Catch Up (Monday) Midterm Exam (Wednesday)		**Midterm (Wednesday)**
Week 9 Oct 15	Logistic regression softmax; multi-way classification	[TSKK] 4.6 [MLaPP] 1.4.6, 8.1-8.3 [ESL] 4.1-4.2, 4.4	
Week 10 Oct 22	neural networks multi-layer perceptron	[Mitchell] Chapter 4 [TSKK] 4.7, 4.8 [MLaPP] 16.5.1-16.5.6, 28 [ESL] 11.3-11.7	
Week 11 Oct 29	Kernel Methods Linear Programming	[MLaPP] 14.1, 14.2.1-14.2.4, 14.4.1, 14.4.3 [ESL] 5.8, 6.3, 6.7 [GT] Chapter 26	
Week 12 Nov 5	Support Vector Machines (SVM) Review: Probability and Statistics	[TSKK] 4.9 [MLaPP] 14.5.2-14.5.4 [ESL] 12.1-12.3	
Week 13 Nov 12	clustering, mixture models, density estimation	[Mitchell] 6.12 [TSKK] 7.1-7.2, 7.5 [MLaPP] 11.1-11.3, 11.4.1-11.4.4, 11.5 [ESL] 14.3.1-14.3.9, 8.5	
Week 14 Nov 19	Computational Learning Theory	[Mitchell] Ch. 7	Topics may change
No Lecture November 21			
Week 15 Nov 26	Genetic Algorithms Learning Sets of Rules Analytical Learning Reinforcement Learning PCA	[Mitchell] Ch. 9 [Mitchell] Ch. 10 [Mitchell] Ch. 11 [Mitchell] Ch. 13 [Bishop] 12.1, 12.2	Topics may change Topics listed are candidates
Friday December 7 : 2-4pm Final Exam			

Recommended Additional References

Abbreviations in brackets is used to reference in the projected course schedule above.

Math Preparation:

- Calculus: Prof. Strang's free online textbook. https://ocw.mit.edu/resources/res-18-001-calculus-online-textbook-spring-2005/
- Linear Algebra: Also from Prof. Strang: https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/index.htm
- [LADW] Linear Algebra Done Wrong by Sergei Treil: https://www.math.brown.edu/~treil/papers/LADW/LADW.html
- [EoLA] "3 Blue 1 Brown" series on Essentials of Linear Algebra: https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab
- Probability and Statistics: Orloff and Bloom: https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/
- The Matrix Cookbook: http://www.math.uwaterloo.ca/~hwolkowi//matrixcookbook.pdf

Additional Machine Learning Resources:

- Max Welling's quick intro: https://www.ics.uci.edu/~welling/teaching/ICS273Afall11/IntroMLBook.pdf
- Alex Smolas's book: http://alex.smola.org/drafts/thebook.pdf
- Gareth James et al's *An Introduction to Statistical Learning* : http://www-bcf.usc.edu/~gareth/ISL/ISLR%20First%20Printing.pdf
- 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
- Andrew Moore's tutorials: https://www.autonlab.org/tutorials
- *Pattern Classification* by Duda, Hart and Stork
- *All of Statistics* by L. Wasserman
- [ESL] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition*
- [MLaPP] Kevin Murphy, *Machine Learning: A Probabilistic Perspective*