**CSE 6243: Advanced topics in machine learning**Instructor: Le Song, School of Computational Science and Engineering

**Course Description**This class will cover several advanced machine learning topics, including graphical models, kernel methods, boosting, bagging, semi-supervised and active learning, and tensor approach to data analysis. The focus of the class will be graphical models and kernel methods, which are currently the major paradigms for building advanced and sophisticated machine learning models for complex real world problems.

Graphical models provide a unified view for a wide range of problems with a very large number of attributes and huge datasets, where we want to obtain a coherent global conclusion from local information. Concepts and algorithms from graphical models enable efficient inference, decision-making and learning in a variety of problems including artificial intelligence, statistics, computer systems, computer vision, natural language processing, and computational biology.

Kernel methods provide a general framework for extending algorithms designed for finding linear relations and patterns to nonlinear cases. Kernel methods approach the problem by mapping the data into a high dimensional feature space; and in that space, a variety of methods can be used to find relations and patterns in the data. Since the mapping can be quite general (eg., not necessarily linear), the relations found in this way are accordingly very general, and the type of data kernel methods can be applied to is also very general (eg., sequence data, audios, images and graph data).

This graduate-level class will provide you with a strong foundation for both applying machine learning to complex real world problems and for addressing core research topics in machine learning. Students entering the class should have a pre-existing working knowledge of probability, statistics, linear algebra and algorithms, though the class has been designed to allow students with a strong numerate background to catch up and fully participate.

**Textbooks**

* [Pattern Recognition and Machine Learning](http://research.microsoft.com/~cmbishop/PRML/index.htm), Chris Bishop
* [Kernel Methods for Pattern Analysis](http://www.kernel-methods.net/), John Shawe-Taylor & Nello Cristianini

Other useful books:

* [Gaussian Processes for Machine Learning](http://www.gaussianprocess.org/gpml/), Carl Edward Rasmussen & Christopher Williams
* [Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond](http://www.amazon.com/Learning-Kernels-Regularization-Optimization-Computation/dp/0262194759) Bernhard Schlkopf & Alex Smola
* [Probabilistic Graphical Models: Principles and Techniques](http://pgm.stanford.edu/) Daphne Koller & Nir Friedman

**Syllabus and Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Lecture & Topics** | **Readings & Useful Links** | **Handouts** |
| Tue 1/10 | Lecture 0: Introduction   * What is Machine Learning? * Applications of Machine Learning * Basic Machine Learning Models * Building Advanced Models with Graphical Models * Building Advanced Models with Kernel Methods * Other Advanced Models * Logistics * Review of probability and conditional independence | [JavaBayes](http://www.pmr.poli.usp.br/ltd/Software/javabayes/)  [Science special issue on data science](http://www.sciencemag.org/site/special/data/)  [Nature special issue on big data](http://www.nature.com/news/specials/bigdata/index.html) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture0.pdf) |
| **Graphical Models** | | | |
| Thu 1/12 | Lecture 1: representation of Bayesian Networks (directed GM)   * Local Markov Assumption encoded by BN * Factorization of distribution according to BN * I-map * D-separation * Limitations of BN | [Bayesian Networks Without Tears](http://scholar.google.com/scholar_url?hl=en&q=http://aaaipress.org/ojs/index.php/aimagazine/article/viewFile/918/836&sa=X&scisig=AAGBfm1cOP1nTgXdktv9JdwH4hQ0SCG02Q&oi=scholarr), E. Charniak, 1991. | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture1.pdf) |
| Tue 1/17 | Lecture 2: representation of undirected GM |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture2.pdf) |
| Thu 1/19 | Lecture 3: unified view of directed and undirected GM |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture3.pdf) |
| Tue 1/24 | Lecture 4: inference --- variable elimination and message passing algorithm |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture4.pdf) |
| Thu 1/26 | Lecture 5: inference --- junction tree algorithm | * [Factor Graph and Sum-Product Algorithm](http://www.comm.utoronto.ca/frank/papers/KFL01.pdf) * [Generalized distribution Law](http://authors.library.caltech.edu/1541/1/AJIieeetit00.pdf) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture5.pdf) |
| Tue 1/31 | Lecture 6: inference --- variational inference | * [Graphical Models, Exponential Families, and Variational Inference](http://www.eecs.berkeley.edu/~wainwrig/Papers/WaiJor08_FTML.pdf) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture6.pdf) |
| Thu 2/2 | Lecture 7: inference --- sampling | * [An Introduction to MCMC for Machine Learning](http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.7133) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture7.pdf) |
| Tue 2/7 | Lecture 8: parameter learning |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture8.pdf) |
| Thu 2/9, Tue 2/14 | Lecture 9, 10: parameter learning from partially observed data --- EM |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture9.pdf) [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture10.pdf) |
| Thu 2/16, Tue 2/21 | Lecture 11, 12: structure learning |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture11.pdf) [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture12.pdf) |
| Thu 2/23 | Lecture 13: Latent Dirichlet Allocation | [Latent Dirichlet Allocation](http://www.cs.princeton.edu/~blei/papers/BleiNgJordan2003.pdf)  [Finding Scientific Topics](http://www.pnas.org/content/101/suppl.1/5228.full.pdf) |  |
| Thu 2/28 | Lecture 14: Kalman Filter, Hidden Markov Models, Conditional Random Fields | [Kalman Filter: Introduction](http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf)  [Hidden Markov Models: Tutorial](http://www.cs.ubc.ca/~murphyk/Bayes/rabiner.pdf)  [Conditional Random Fields](http://www.cis.upenn.edu/~pereira/papers/crf.pdf)  [Conditional Random Fields: Introduction](http://www.inference.phy.cam.ac.uk/hmw26/papers/crf_intro.pdf) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture13.pdf) |
| Tue 3/1 | Lecture 15: Collaborative Filtering | [Probablistic Matrix Factorization](http://www.cs.utoronto.ca/~amnih/papers/pmf.pdf)  [Probabilistic Matrix Factorization using MCMC](http://www.cs.utoronto.ca/~amnih/papers/bpmf.pdf) | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture14.pdf) |
| **Kernel Methods** | | | |
| Thu 3/8 | Lecture 16: kernels, kernel classifier and regression |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture15.pdf) |
| Tue 3/13 | Lecture 17: kernel PCA, clustering, canonical correlation analysis |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture16.pdf) |
| Thu 3/15 | Lecture 18: two sample test and measure of dependence beyound vector data (eg. sequences and graphs) |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture17.pdf) |
| Tue 3/27 | Lecture 19: Gaussian process I |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture18.pdf) |
| Thu 3/29 | Lecture 20: Gaussian process II |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture19.pdf) |
| Tue 4/3 | Lecture 21: fast kernel methods and random features |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture20.pdf) |
| **Other Advanced Models and Topics** | | | |
| Thu 4/5 | Lecture 22: bagging and boosting |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture21.pdf) |
| Tue 4/10 | Lecture 23: semi-supervised learning |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture22.pdf) |
| Thu 4/12 | Lecture 24: active learning |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture23.pdf) |
| Tue 4/17 | Tensor data analysis I |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture_tensors1.pdf) |
| Thu 4/19 | Tensor data analysis II |  | [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/lecture_tensors2.pdf) [Slides](http://www.cc.gatech.edu/~lsong/teaching/8803ML/review.pdf) |
| **Project Presentations** | | | |
| Tue 4/24 | Project Presentation |  |  |
| Thu 4/26 | Project Presentation |  |  |