

CG168: Human and Machine Learning

Mark Johnson

Fall 2008

This course has two main goals:

- to introduce you to the ideas and techniques of machine learning and the probabilistic models that lie behind them, and
- to apply these to a project involving natural language processing, specifically learning to parse with dependency grammars.

Roughly speaking, machine learning techniques divide into two main kinds. *Supervised learning techniques* are useful when you have labeled training data, i.e., examples of the kind of output you want the system to provide. In our project we'll use supervised learning techniques to learn how to parse novel sentences given example parses of other sentences. *Unsupervised learning techniques* are useful when you have unlabeled training data; later in our project we'll learn how to parse from training data that consists of sentences alone.

1 Administrative details

Instructor:

Mark Johnson (Professor)
Cognitive and Linguistic Sciences, Box 1978
Metcalf Research Building, Room 231
190 Thayer Street
Phone: 863 1670, email: Mark_Johnson@Brown.edu
Office Hours: Wednesday 10am-noon.

Teaching assistant:

William Headden
email: headden@gmail.com

Class time and location:

CIT 367, Tuesday and Thursday 9:00–10:30am

2 Grading and textbooks

Half of grade will be based on the in-class mid-term and final exams, and half based on the project. You can get 200-level credit by doing extra work in the class project.

Recommended textbook: Bishop (2006) *Pattern Recognition and Machine Learning*.

3 Syllabus

Week 1: Administrative details. Course and project overview.

Week 2: Dependency parsing: Dependency grammar and incremental dependency parsing. Project goal: map dependency parses to sequences of parser actions and vice versa.

Week 3: Supervised learning: K-nearest neighbour and Naive Bayes classifiers. Curse of dimensionality. Bias-variance dilemma. Project goal: learn a Naive Bayes classifier that predicts best parser action given current state.

Week 4: Supervised learning: Maximum entropy classifiers. Stochastic gradient ascent and the Perceptron. Regularization. 200-level project goal: learn a Maximum Entropy classifier that predicts best parser action given current state.

Week 5: Supervised learning: Structural risk minimization and Support Vector Machines. Kernel methods.

Week 6: Dependency parsing: non-deterministic parsing and search. Particle filtering. 200-level project goal: implement a particle-filtering non-deterministic dependency parser.

Week 7: Dependency parsing: dynamic programming algorithms for finding most likely parse and sampling parses. Project goal: implement a dynamic programming Viterbi parser for dependency grammars. 200-level project goal: implement a sampler for dependency grammars.

Week 8: Unsupervised learning: K-means clustering. Agglomerative and divisive clustering algorithms.

Week 9: Unsupervised learning: Expectation Maximization (EM). Dynamic programming algorithms for calculating dependency expectations. Project goal: learn a dependency grammar from strings alone using EM.

Week 10: Unsupervised learning: Variational Bayesian (VB) inference. Project goal: modify EM dependency grammar learning procedure to use VB instead.

Week 11: Unsupervised learning: Markov Chain Monte Carlo (MCMC) inference. Gibbs sampling.

Week 12: Unsupervised learning: MCMC inference procedures. Slice sampling. Metropolis-Hastings sampling. 200-level project goal: implement a dependency grammar learner using a Metropolis-within-Gibbs sampler.

Week 13: Unsupervised learning: Importance sampling. Particle filtering. 200-level project goal: implement a dependency grammar learner using a particle filter.

Week 14: Unsupervised learning: Co-training and self-training.