Introduction to Machine Learning

(CMPUT 466/551), University of Alberta, Edmonton, Canada, 2009 Fall
Instructors: Russell Greiner and Barnabás Póczos

Learning—i.e., using experience to improve performance—is an essential component of intelligence. The field of Machine Learning, which addresses the challenge of producing machines that can learn, has become an extremely active and exciting area with an ever expanding inventory of practical (and profitable!) results, many enabled by recent advances in the underlying theory. This course provides an introduction to the field with an emphasis on the design of agents that can learn about their environment to help them improve their performance on a range of tasks. We will cover practical aspects including algorithms for learning predictors, various linear models, separators, regressors, decision trees, neural networks, support vector machines, and belief networks. We will discuss supervised and unsupervised learning, as well as theoretical concepts, including relevant ideas from statistics, Bayesian learning, and the PAC learning framework.

Foundations of Probability

Bayes rule. Independence. Moments. MLE estimation. MAP estimation. Gaussian distribution.

Learning Linear Models

Learning linear regression models.
General issues for predictors: evaluating predictors, model selection: bias-variance, regularization.
Linear classifiers : perceptron, logistic regression, linear discriminant analysis.
(Linear) support vector machines: duality, Lagrange method, quadratic programming.

Learning Non-Linear Models

Kernel foundations. Learning kernel classifiers. Support vector machines.
Artificial neural nets: backpropagation, line search, conjugate gradient.
Decision trees: entropy, pruning.
Gaussian processes.

Ensemble Methods

Bagging. Boosting.

Computational Learning Theory

PAC learning. VC-dimension.

Graphical Models

Directed models: belief nets, parameter estimation.
Dynamic Bayesian net: hidden Markov models, linear dynamical systems.
Learning Bayesian net structure.
Undirected models: Markov random fields.

Unsupervised Learning

Principle components analysis.
Independent component analysis. Independent subspace analysis. Independent process analysis.
Nonnegative matrix factorization.
K-means clustering. Hierarchical clustering.