Machine Learning: Other readings

CS229 covered a broad swath of topics in machine learning, compressed into a single quarter. Machine learning is a large but still growing field, with thousands of new research papers written each year. The "Related AI classes" handout posted on the course website describes some classes that you can take to learn more about AI and Machine Learning. In addition, here are some sources from which you can read more about this fascinating topic.

General machine learning

If you are generally interested in reading more about machine learning, but not necessarily on any specific topic, a good starting place (focusing on supervised and unsupervised learning) would be:

- **Pattern Recognition and Machine Learning**. Christopher Bishop. Springer, 2007.

In addition, if you are interested in learning more about a specific topic covered in CS229, below we describe some excellent resources that focus on different subsets of machine learning.

Generalized linear models

The following is a "classic" reference for GLMs, and goes into exponential family distributions and different versions of GLMs in great detail:

- **Generalized Linear Models**. P. McCullagh and J.A. Nelder. Chapman & Hall, 1989.

SVMs and kernels

This describes in detail the math behind SVMs (including convexity and duality), kernels, and also talks about different variations on the SVM algorithm you learned in CS229:

- **Support Vector Machines and other Kernel-based Learning Methods**. John Shawe-Taylor & Nello Cristianini. Cambridge University Press, 2000.

Independent components analysis

The following book derives a few variations of and explanations for PCA, outlines different ways of deriving ICA algorithms, and gives many examples.

- **Independent Component Analysis**. Aapo Hyvrinen, Juha Karhunen and Erkki Oja. Wiley-Interscience. 2001.

Learning Theory

- **Lecture notes by Sham Kakade** at TTI/University of Chicago, on learning theory. Available at http://ttic.uchicago.edu/ tewari/LT_SP2008.html
- **Neural Network Learning: Theoretical Foundations**. Martin Anthony and Peter L. Bartlett. Cambridge University Press. 1999.

Reinforcement learning and MDPs

Of the two sets of books below, the first is more mathematical, and rigorously discusses the math behind MDPs; the latter covers a broad range of the concepts encountered in coming up with reinforcement learning algorithms and using them.

- **Dynamic Programming and Optimal Control (Vol I & II)**. Dimitri P. Bertsekas. Athena Scientific. 2000.
- **Reinforcement Learning: An Introduction**. Richard Sutton and Andrew Barto. MIT Press 1998.

Other sources

Finally, here are a number of other excellent textbooks on machine learning, AI, and related topics:

- **Pattern Classification, 2nd ed**. Richard Duda, Peter Hart and David Stork. 2001.
- **Artificial Intelligence: A Modern Approach (2nd ed)**. Stuart Russell and Peter Norvig. Prentice Hall. 2002.
- **Introduction to Computational Learning Theory**. Michael J. Kearns and Umesh V. Vazirani. MIT Press. 1994.
- **Probabilistic Graphical Models: Principles and Techniques**. Daphne Koller and Nir Friedman. MIT Press. To be published.
- **Machine Learning**. Tom Mitchell. McGraw-Hill, 1997.

Machine learning is a rapidly advancing field, and so many of the concepts you learned in CS229 reflect relatively recent developments, that have not yet made it into textbooks. To learn about the most recent advances at the forefront of machine learning, take a look at the last few years of conference proceedings of ICML (International Conference on Machine Learning) and NIPS (Neuro Information Processing Systems). In addition, some other related conferences that have many machine learning papers include UAI (Uncertainty in AI), IJCAI (International Joint conference on AI) and AAAI (Association for the Advancement of AI). Finally, a range of conferences in NLP, robotics, computer vision, computational biology, medical imaging, and *many* other fields also routinely contain papers that describe applications of learning to these different areas.