Useful Resources

Below is a compilation of web links. Hopefully these resources will help improve your learning experience.

Informative Web Sites

- Java Applets for Machine Learning Note: The applets are in German ** Page not accesible. Message: The page you want to visit cannot be displayed.
- A Brief Introduction to Machine Learning by Gunnar Ratsch
- CS229 Machine Learning - Stanford - This is the Stanford CS course on Machine Learning that Prof Ng has taught for a number of years. The material parallels the Coursera course, but covers some additional topics and goes into much more depth on the mathematics.
- Dive into Machine Learning compiles a variety of resources, taking a hack-first approach so you can get "hooked." Prof. Ng's course is the centerpiece.
- Cornell Virtual Workshop Training on programming languages, parallel computing, code improvement, and data analysis.

Linear Algebra

- Introduction to Linear Algebra
- CS 229 Section notes on Linear Algebra
- Free linear algebra book with solutions

Writing Equations in Forum Posts

- Short Guide to LaTex Math Here is a quick guide to entering equations using LaTeX. The directives are inserted between two dollar signs, $Forexample, thefractionforonehalfisenteredas\frac{1}{2}$$, and displays as 12.
- LaTex Math Tutorial

Online E-Books

- Introduction to Machine Learning by Nils J. Nilsson

- Introduction to Machine Learning by Alex Smola and S.V.N. Vishwanathan
- Introduction to Data Science by Jeffrey Stanton The link appears to be dead, here is another.
- Bayesian Reasoning and Machine Learning by David Barber
- Understanding Machine Learning, © 2014 by Shai Shalev-Shwartz and Shai Ben-David
- Elements of Statistical Learning, by Hastie, Tibshirani, and Friedman
- Pattern Recognition and Machine Learning, by Christopher M. Bishop

Textbook information

- (none)

Advanced classes online

- Andrew Ng's advanced lectures - YouTube
- Machine Learning - CosmoLearning
- Machine Learning - AcademicEarth
- Learning from Data - Caltech
- Machine Learning - MIT
- Machine Learning - U. of Washington - via Coursera
- Big Data, Large Scale Machine Learning - NYU (not a MOOC)
- Machine Learning UBC 2013 - Youtube
- Neural Networks Demystified

Machine Learning frameworks and libraries in Python

- PyBrain: Various machine learning algorithms for Python programmers. Focuses on neural networks.
- PyML: Machine Learning object oriented framework for Linux and Mac OS X focused on classification and regression by Asa Ben-Hur.
- scikit-learn: Comprehensive Machine Learning toolkit for Python (based on SciPy with numpy and mathplotlib). "Ipython -pylab" provides interactive environment like Octave - scikit-learn provides optimized implementations of pretty well everything (using fast libraries like liblinear and libsvm). Should be used instead of Octave for research prototyping, production and especially for education.

Machine Learning frameworks and libraries in C++

- mlpack: a scalable C++ machine learning library.
- SHARK: a fast, modular, feature-rich open-source C++ machine learning library.
- Dlib-ml: A Machine Learning Toolkit.
- Waffles: A collection of command-line tools for researchers in machine learning, data mining, and related fields. All of the functionality is also provided in a clean C++ class library.
- MLC++: a library of C++ classes for supervised machine learning.

Machine Learning frameworks and libraries in Java

- Weka: A collection of machine learning algorithms for data mining tasks.
- Apache Mahout: A scalable machine learning library .
- LIBLINEAR : LIBLINEAR -- A Library for Large Linear Classification. I think this link was mentioned in one of the lectures.
- Deeplearning4j: Open-source, distributed, deep-learning library for the JVM. Integrated with Hadoop and Spark, DL4J is designed to be used on distributed GPUs and CPUs.

Machine Learning Data Sets

- Links to many ML data repositories
- UCI Machine Learning Repository - Univ of California Irvine
- Kaggle: Machine Learning and data mining activities
- COCO-Text: Dataset for Text Detection and Recognition

Octave packages

- http://octave.sourceforge.net/ GNU Octave packages development and repository.

Octave online

- http://octave-online.net/

Translation Projects

- Mexico Study Group Notes

Useful papers

- Massive collection of academic papers are available here: Machine Learning Library.

General

- Domingos, Pedro. "A few useful things to know about machine learning." Communications of the ACM 55, no. 10 (2012): 78-87
- Shewchuk, Jonathan Richard. "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain." 1994
- To understand cost functions better An Introduction To Understanding Cost Functions

Boosting

- Friedman, J. H. "Greedy Function Approximation: A Gradient Boosting Machine." (Feb. 1999a)
- Ridgeway, Greg. "Generalized Boosted Models: A guide to the gbm package." Update 1 (2007): 1.
- Rojas, Raúl. "AdaBoost and the Super Bowl of Classifiers A Tutorial Introduction to Adaptive Boosting." Freie University, Berlin (2009).

Outlier and Anomaly Detection

- Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Outlier detection: A survey." ACM Computing Surveys, to appear (2007).
- Kriegel, Hans-Peter, Peer Kröger, and Arthur Zimek. "Outlier detection techniques." In Tutorial at the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining. 2009.

SVM

- "An Idiot's Guide to Support Vector Machines"

http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf

Interesting applications

- Castillo, Carlos, Marcelo Mendoza, and Barbara Poblete. "Information credibility on Twitter." In Proceedings of the 20th international conference on World wide web, pp. 675-684. ACM, 2011.
- Norman, Kenneth A., Sean M. Polyn, Greg J. Detre, and James V. Haxby. "Beyond mind-reading: multi-voxel pattern analysis of fMRI data."Trends in cognitive sciences 10, no. 9 (2006): 424-430.
- Pereira, Francisco, Tom Mitchell, and Matthew Botvinick. "Machine learning classifiers and fMRI: a tutorial overview." Neuroimage 45, no. 1 Suppl (2009): S199.
- Dean Pomerleau Autonomous Driving (link)