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](http://www.ml.inf.ethz.ch/education/lectures_and_seminars/annex_estat/index.html) 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](http://events.ccc.de/congress/2004/fahrplan/files/105-machine-learning-paper.pdf)
* [CS229 Machine Learning - Stanford](http://cs229.stanford.edu/materials.html) - 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](https://github.com/hangtwenty/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](https://cvw.cac.cornell.edu/default) Training on programming languages, parallel computing, code improvement, and data analysis.

Linear Algebra

* [Introduction to Linear Algebra](http://www.eigenvector.com/Docs/LinAlg.pdf)
* [CS 229 Section notes on Linear Algebra](http://cs229.stanford.edu/section/cs229-linalg.pdf)
* [Free linear algebra book with solutions](http://joshua.smcvt.edu/linearalgebra/)

Writing Equations in Forum Posts

* [Short Guide to LaTex Math](ftp://ftp.ams.org/pub/tex/doc/amsmath/short-math-guide.pdf) Here is a quick guide to entering equations using LaTeX. The directives are inserted between two dollar signs. For example, the fraction for one half is entered as \$\$ \frac{1}{2}\$\$ ,(without any escapes before the dollar signs) and displays as 12.
* [LaTex Math Tutorial](http://www.forkosh.com/mimetextutorial.html)

Online E-Books

* [Introduction to Machine Learning by Nils J. Nilsson](http://robotics.stanford.edu/~nilsson/MLBOOK.pdf)
* [Introduction to Machine Learning by Alex Smola and S.V.N. Vishwanathan](http://alex.smola.org/drafts/thebook.pdf)
* [Introduction to Data Science by Jeffrey Stanton](http://jsresearch.net/wiki/projects/teachdatascience/Teach_Data_Science.html) The link appears to be dead, [here is another](http://surface.syr.edu/cgi/viewcontent.cgi?article=1165&context=istpub).
* [Bayesian Reasoning and Machine Learning by David Barber](http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.Online)
* [Understanding Machine Learning, © 2014 by Shai Shalev-Shwartz and Shai Ben-David](http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/copy.html)
* [Elements of Statistical Learning, by Hastie, Tibshirani, and Friedman](http://statweb.stanford.edu/~tibs/ElemStatLearn/)
* [Pattern Recognition and Machine Learning, by Christopher M. Bishop](http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%202006.pdf)

Textbook information

* (none)

Advanced classes online

* [Andrew Ng's advanced lectures - YouTube](http://www.youtube.com/course?list=ECA89DCFA6ADACE599)
* [Machine Learning - CosmoLearning](http://www.cosmolearning.com/courses/machine-learning/)
* [Machine Learning - AcademicEarth](http://www.academicearth.org/courses/machine-learning/)
* [Learning from Data - Caltech](http://work.caltech.edu/previous.html)
* [Machine Learning - MIT](http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-867-machine-learning-fall-2006/)
* [Machine Learning - U. of Washington - via Coursera](https://www.coursera.org/course/machlearning)
* [Big Data, Large Scale Machine Learning - NYU (not a MOOC)](http://cilvr.cs.nyu.edu/doku.php?id=courses:bigdata:start)
* [Machine Learning UBC 2013 - Youtube](https://www.youtube.com/playlist?list=PLE6Wd9FR--EdyJ5lbFl8UuGjecvVw66F6)
* [Neural Networks Demystified](https://www.youtube.com/playlist?list=PLiaHhY2iBX9hdHaRr6b7XevZtgZRa1PoU)

Machine Learning frameworks and libraries in Python

* [PyBrain](http://pybrain.org/): Various machine learning algorithms for Python programmers. Focuses on neural networks.
* [PyML](http://pyml.sourceforge.net/): Machine Learning object oriented framework for Linux and Mac OS X focused on classification and regression by Asa Ben-Hur.
* [scikit-learn](http://scikit-learn.org/stable/): 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.
* [tensor-flow](https://www.tensorflow.org/): open source software library for machine learning.

Machine Learning frameworks and libraries in C++

* [mlpack](http://www.mlpack.org/): a scalable C++ machine learning library.
* [SHARK](http://image.diku.dk/shark/sphinx_pages/build/html/index.html): a fast, modular, feature-rich open-source C++ machine learning library.
* [Dlib-ml](http://dlib.net/ml.html): A Machine Learning Toolkit.
* [Waffles](http://waffles.sourceforge.net/): 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++](http://www.sgi.com/tech/mlc/): a library of C++ classes for supervised machine learning.

Machine Learning frameworks and libraries in Java

* [Weka](http://www.cs.waikato.ac.nz/ml/weka/): A collection of machine learning algorithms for data mining tasks.
* [Apache Mahout](http://mahout.apache.org/): A scalable machine learning library .
* [LIBLINEAR](http://www.csie.ntu.edu.tw/~cjlin/liblinear/): LIBLINEAR -- A Library for Large Linear Classification. I think this link was mentioned in one of the lectures.
* [Deeplearning4j](http://deeplearning4j.org/): 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](http://www.kdnuggets.com/datasets/)
* [UCI Machine Learning Repository - Univ of California Irvine](http://archive.ics.uci.edu/ml/)
* [Kaggle: Machine Learning and data mining activities](http://www.kaggle.com/)
* [COCO-Text: Dataset for Text Detection and Recognition](http://vision.cornell.edu/se3/coco-text/)

Octave packages

* <http://octave.sourceforge.net/> GNU Octave packages development and repository.

Octave online

* <http://octave-online.net/>

Translation Projects

* [Mexico Study Group Notes](https://share.coursera.org/wiki/index.php?title=ML:spanishMain)

Useful papers

* Massive collection of academic papers are available here: [Machine Learning Library](http://www.bradblock.com.s3-website-us-west-1.amazonaws.com/mll.html).

General

* Domingos, Pedro. ["A few useful things to know about machine learning."](http://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf) Communications of the ACM 55, no. 10 (2012): 78-87
* Shewchuk, Jonathan Richard. ["An Introduction to the Conjugate Gradient Method Without the Agonizing Pain."](http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf) 1994
* To understand cost functions better [An Introduction To Understanding Cost Functions](https://www.youtube.com/watch?v=euhATa4wgzo)

Boosting

* Friedman, J. H. ["Greedy Function Approximation: A Gradient Boosting Machine."](http://docs.salford-systems.com/GreedyFuncApproxSS.pdf) (Feb. 1999a)
* Ridgeway, Greg. ["Generalized Boosted Models: A guide to the gbm package."](http://gradientboostedmodels.googlecode.com/git-history/c532a997943c634ead7e27cdf54dbe343c2c7110/svn/pkg/inst/doc/gbm.pdf) Update 1 (2007): 1.
* Rojas, Raúl. ["AdaBoost and the Super Bowl of Classifiers A Tutorial Introduction to Adaptive Boosting."](http://www.inf.fu-berlin.de/inst/ag-ki/adaboost4.pdf) Freie University, Berlin (2009).

Outlier and Anomaly Detection

* Chandola, Varun, Arindam Banerjee, and Vipin Kumar. ["Outlier detection: A survey."](http://www.bradblock.com.s3-website-us-west-1.amazonaws.com/Outlier_Detection_A_Survey.pdf) ACM Computing Surveys, to appear (2007).
* Kriegel, Hans-Peter, Peer Kröger, and Arthur Zimek. ["Outlier detection techniques."](http://www.dbs.ifi.lmu.de/~zimek/publications/KDD2010/kdd10-outlier-tutorial.pdf) 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."](http://www.ra.ethz.ch/cdstore/www2011/proceedings/p675.pdf) 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."](http://www.cs.princeton.edu/courses/archive/spr07/cos424/papers/NormanEtAlTICS.pdf)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."](http://www.stanford.edu/group/bad/talks/classification_fmri.pdf) Neuroimage 45, no. 1 Suppl (2009): S199.
* Dean Pomerleau Autonomous Driving [(link)](http://www.ri.cmu.edu/research_project_detail.html?type=publication&project_id=160&menu_id=261)