Machine Learning
Resource Guide

Jason Brownlee

MACHINE
LEARNING
MASTERY

Machine Learning Mastery
Web: http://MachineLearningMastery.com
Email: jason@MachineLearningMastery.com

Machine Learning Resource Guide
by Jason Brownlee, PhD

Share this Guide
If you know someone who can benefit from this guide, just send them this link:

http://MachineLearningMastery.com/machine-learning-resources/

Table of Contents

Introduction

Hi there, my name is Jason from MachineLearningMastery.com. Thanks for downloading this ***Machine Learning Resource Guide***.

I have worked hard to collect and list only the best resources that will help you jump-start your journey towards machine learning mastery. I've categorized the resources into main themes such as videos, books and courses.

I'm certain you will find great value in the resources listed in this guide. Take your time and select a medium or resource type you prefer and start working through resources one-by-one. Try not to overload yourself. Remember to think hard about what you want from a resource and actively take notes.

I'm interested to hear what resources you try, send me an email and let me know via jason@MachineLearningMastery.com or visit my site MachineLearningMastery.com and leave a comment. I hope to hear from you soon.

Jason Brownlee.

Books

I read a lot of books, and even in this age of ebooks, I like having a lot of reference books on the bookshelf. I also like having books in PDF form so I can search them quickly and pull out the information I need.

The books listed in this section are grouped by a few different criteria that you may find useful, as such you may see a few duplicates across the different lists of books.

I have provided links to each book on Amazon. The links are affiliate links which means I will get a few cents from Amazon if you decide to buy a book.

My advice: Pick one book and read it, cover-to-cover.

Beginner Books

These are books for the absolute beginner to get a feeling for what machine learning or working with data is all about. From a business and semi-technical perspective.

- Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die
- Data Science for Business: What you need to know about data mining and data-analytic thinking
- Data Smart: Using Data Science to Transform Information into Insight

Practical Books

If you are a programmer or engineer and are looking for a book with code examples to implement or execute, these are books for you:

- Programming Collective Intelligence: Building Smart Web 2.0 Applications
- Data Mining: Practical Machine Learning Tools and Techniques
- Machine Learning for Hackers
- Machine Learning: An Algorithmic Perspective
- Machine Learning in Action
- Applied Predictive Modeling

You can learn more about these books in my blog post 6 Practical Books for Beginning Machine Learning

Python Books

These are books for learning and applying machine learning if you are a python programmer.

- Building Machine Learning Systems with Python
- Learning scikit-learn: Machine Learning in Python
- Machine Learning in Action

- Programming Collective Intelligence: Building Smart Web 2.0 Applications
- Machine Learning: An Algorithmic Perspective
- Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More
- Natural Language Processing with Python
- Programming Computer Vision with Python: Tools and algorithms for analyzing images
- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython

You can learn more about these books in my blog post: Python Machine Learning Books

R Books

If you are an R programmer or are looking at applying machine learning in R, these books are for you.

- Applied Predictive Modeling
- An Introduction to Statistical Learning: with Applications in R
- Practical Data Science with R
- Machine Learning with R
- Data Mining with R: Learning with Case Studies
- Data Mining and Business Analytics with R
- Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery

You can learn more about these books in my blog post: Books for Machine Learning with R

Textbooks

These are books for machine learning practitioners looking to go beyond the practical books and deeper into theory. These are textbooks commonly used in undergraduate and postgraduate university courses.

- Machine Learning, by Tom Mitchell
- Learning From Data, by Yaser Abu-Mostafa, Malik Magdon-Ismail and Hsuan-Tien Lin
- Machine Learning: A Probabilistic Perspective, by Kevin Murphy
- Pattern Recognition and Machine Learning, by Christopher Bishop
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani and Jerome Friedman

Communities

You will have a lot of questions along your journey toward machine learning mastery and there are excellent places where machine learning experts can answer those questions for you, if you know where to look.

Each site listed below allows you to create an account for free and ask your question. Review the types of questions and answers offered in each community before selecting the right community for you to ask your question.

It is very likely your question has been asked and answered before. Try searching for it on each community before posting.

Stack Exchange

The stack exchange sites are question and answer communities, so they are targeted towards problem solving. You can post the specific questions you have, answer questions to which you know the answer and (my favorite) read questions and answers to discover new methods and perspectives.

There are four sites I like to dip into:

- [Cross Validated](): This site is useful for low-level questions on algorithms and statistical methods.
- [Quantitative Finance](): (specifically the [machine learning tag]()) This site is useful if you are operating in the financial domain, but generally if you are working with time series data.
- [Programmers](): (specifically the [machine learning tag]()) Great for specific code questions, such as a problem with a given library or tool you are using.
- [Stack Overflow](): (specifically the [machine learning tag]()) Again, like programmers, great for specific questions with the implementation side of machine learning. It's also the oldest site and can cover machine learning algorithms and libraries.

There is a new site that has started up, but is still in beta, so it may not survive. It is called [Data Science]() and I am finding it very interesting for the general concerns of applied machine learning (mix of code and math).

Reddit

Reddit is a community of communities called sub-reddits. A given subreddit can be question and answer site, a link sharing site or (more typically) a mix of the two.

A few subreddits I frequent include:

- [Machine Learning](): Contains of mix of "how do I get started" and more advanced links to machine learning blog posts. Also good for linking to your own projects to get some feedback.

- **Computer Vision**: Mostly questions on computer vision questions both theoretical and practical (such as libraries).
- **Natural Language**: Focus on natural language processing, providing a good mix of questions and links to relevant articles and blog posts.
- **Statistics**: Discussion on statistical software and methods, great for digging deeper into a given method or algorithm.
- **Data Science**: Mostly links to posts that straddle data analysis and machine learning.
- **Big Data**: Focused posts and discussions on the big data ecosystem.

There are other sub-reddits on relevant and related topics, but I have not found them as useful.

Quora

Quora is a question and answer site that is divided into topics, much like reddit but only questions and answers. The questions are typically good and the answers high quality. Unlike the stack exchange sites, they are typically less technical, less problem focused and more meaty.

A few Quora topics I frequent include:

- **Machine Learning**: Useful for high-level questions on algorithms, processes, resources and getting started. A good mix.
- **Statistics**: Focus on deeper statistical methods and algorithms, but includes a lot of machine learning content.
- **Data Mining**: Good questions with a focus on the applied side of machine learning, but a lot of overlap with Machine Learning.
- **Data Science**: Much like the Data Mining and Machine learning topics, the questions are typically a higher level.

There are many other topics that might be useful, not limited to Data Analysis, Predictive Analytics, NLP and Computer Vision. Also there are topics on specific methods such as SVM, Deep Learning, Classification, and R.

Other

There are some other great communities around that I could not classify as easily.

- **MetaOptimize Q+A**: Like Cross Validated, this is a question and answer site that is great for lower level questions on specific algorithms and methods. Maths and theory heavy.
- **Kaggle Forums**: Great for discussion around specific competitions and datasets, and full of great nuggets of advice for feature engineering, ensembling and refining your test harnesses.
- **Data Tau**: A social news site with a focus on links to posts on data and machine learning relate topics. Low traffic and useful links.

Videos

Videos are a great way to learn about machine learning, both for lecture and tutorial content.

University Courses

There are university courses that are offered online for free by organizations such as Coursera and edX. They include video lectures, homework that is assessed, quizzes and tests. Some of the courses also have just the video lectures listed on sites like YouTube. Try searching.

- Stanford: Machine Learning, by Andrew Ng
- Stanford: Probabilistic Graphical Models, by Daphne Koller
- Caltech: Learning from Data, by Yaser Abu-Mostafa
- University of Toronto: Neural Networks for Machine Learning, by Geoffrey Hinton
- University of Washington: Machine Learning, by Pedro Domingos
- University of Washington: Introduction to Recommender Systems, by Joseph Konstan and Michael Ekstrand
- University of Washington: Introduction to Data Science, by Bill Howe

Paid Courses

There are paid courses on Machine Learning offered by organizations such as Udemy. You pay a fee and have access to the premium content to learn something specific.

- Udemy: An Intro to Machine Learning with Web Data, by Hilary Mason
- Udemy: Advanced Machine Learning, by Hilary Mason
- Udemy: Introduction to R, by Jagannath Rajagopal
- Udemy: Working with Big Data, by Pearson

Other Videos

- Machine Learning Category on VideoLectures.Net
- "Getting In Shape For The Sport Of Data Science" - Talk by Jeremy Howard
- Facebook Tech Talk: Peter Norvig on big data

University Course Material

There is a popular trend for top-level technical universities to put course materials online including lecture videos, slides, homework and assignments. This material can be used for self-study.

Some universities make the materials easier to find than others, MIT is a shining light in this regard with their OpenCourseWare initiative.

Undergraduate Level

- [MIT 6.034 Artificial Intelligence](#) (provide a machine learning focus)
- [MIT 15.075 Statistical Thinking and Data Analysis](#)
- [Stanford CS229 Machine Learning](#) ([SEE site](#))
- [Stanford Statistics 315a Modern Applied Statistics: Elements of Statistical Learning](#)
- [Stanford Statistics 315b Modern Applied Statistics: Elements of Statistical Learning II](#)
- [Caltech Learning from Data](#)

Graduate Level

- [MIT 6.867 Machine Learning](#)
- [MIT 6.825 Techniques in Artificial Intelligence](#) (related machine learning topics)
- [MIT 9.520 Statistical Learning Theory and Applications](#)
- [MIT 9.641 Introduction to Neural Networks](#)
- [MIT 15.097 Prediction: Machine Learning and Statistics](#)
- [MIT 18.465 Topics in Statistics: Statistical Learning Theory](#)
- [Harvard CS281 Intelligent Machines: Perception, Learning, and Uncertainty](#) (also [CS181](#))
- [Cornell CS6784 Advanced Machine Learning](#)
- [CMU 10-701 Machine Learning](#)(videos [here](#) and [here](#))

Software and Libraries

There are a lot of software and libraries that you could use for machine learning.

Below are some best-of-breed software tools and libraries that are useful for learning and practicing machine learning.

- WEKA (GUI, Java)
- R
- Scikit-Learn (Python)
- Octave (an open source MatLab)
- BigML (in the browser)

If you are just starting out, I recommend using the Weka graphical user interface. For example, you can run your first classifier in 5 minutes flat.

If you are struggling with which programming language to use, check out my post: Best Programming Language for Machine Learning.

If you are a Java programmer you may be interested in my post: Java Machine Learning.

Competitions

Competitions are common with Artificial Intelligence and Machine Learning conferences. Take a look at the webpages for some of the popular conferences and you will very likely find current active machine learning competitions.

Competitive machine learning can be a great way to learn new data preparation and modelling techniques. People in and around competitive machine learning can provide a wealth of tips, resources and different ways of approaching the same problem. The competitions can also be a great way to test out methods and ideas.

- [Kaggle](#)
- [TunedIT](#)
- [CrowdAnalytix](#)
- [InnoCentive](#)
- [Challenge.gov](#)
- [KDDCup](#)

Guides

I have a passion for helping programmers and engineers get started and kick ass with machine learning. You can shortcut your machine learning journey with a number of the guides and courses that I have created for you.

Beginner

- [Self-Study Guide to Machine Learning](#): (**Start Here!**) Discover the structured framework for self-studying machine learning that includes 4 competency levels and focused objectives and activities for each level.
- [Machine Learning Foundations](#): Discover the concepts and definitions of machine learning and have the confidence to explain it to friends and colleagues.
- [Conquer Self-Limiting Beliefs in Machine Learning](#): Discover your own self-limiting beliefs that are halting you from getting started or making progress in the field of machine learning.
- [Machine Learning Matters](#): Discover why machine learning matters to you and why it matters to the world.

Novice

- [Applied Machine Learning Process](#): Discover the structured step-by-step process for applying machine learning to your own problems now and in the future. Includes a clear 6-step framework with activities and questions to answer at each step along the way.
- [Jump-Start Scikit-Learn](#): Discover the Python machine learning library scikit-learn in this lightweight recipe book. Contains 35 recipes ready to copy and paste for data handling, supervised learning, regularization algorithms, ensemble methods and advanced topics.
- [Jump-Start Weka](#): Discover the Weka machine learning workbench including step-by-step tutorials for analyzing data, applying machine learning algorithms and designing and interpreting machine learning experiments.
- [Beginning Weka](#): [**Video Course**] Discover the process of applied machine learning with step-by-step tutorials and worked case study problems using the Weka machine learning workbench. The feature of this course are the 3 real-world case studies with step-by-step tutorials and videos.

Intermediate

- [Small Projects Methodology](#): Discover the blueprint for learning and practicing applied machine learning with 4 project types and 90 project ideas.
- [Algorithm Description Template](#): Discover a strategy for learning a machine learning algorithm fast. I used this strategy to learn and describe 45 nature inspired algorithms that I turned into a book.
- [Clever Algorithms: Nature-Inspired Programming Recipes](#): Discover 45 nature-inspired algorithms described consistently using a structured algorithm template. All algorithms include a working implementation in Ruby.

Jet Fuel

- [Super Bundle](http://MachineLearningMastery.com): [**Get It All!**] In this bundle you get a copy of all current and all future standalone machine learning mastery guides and courses. As a part of this super bundle you will be emailed as new guides are added in the future so that you can download them at no extra cost.

Connect With Me!

Hey, my name is Jason. I'm from Australia and I have a Masters and Phd in Artificial Intelligence, I've written books on algorithms, consulted for startups and I work on tropical cyclone forecasting systems. I get a lot of satisfaction helping programmers make their start and kick some ass with machine learning.

I am 33 years old, married with a young son and in my free time I like to read books, code, write articles and participate in machine learning competitions.

You can learn more about me and my story by clicking here.

Reach out to me, I'd love to hear from you and your goals with machine learning.

Contact me via email on jason@MachineLearningMastery.com

Follow me on:

LinkedIn: https://www.linkedin.com/in/jasonbrownlee
Twitter: http://twitter.com/TeachTheMachine
Facebook: http://www.facebook.com/pages/Machine-Learning-Mastery/1429846323896563
Google+:   http://plus.google.com/+MachineLearningMasteryHome