

05.15-Learning-More

June 27, 2018

This notebook contains an excerpt from the [Python Data Science Handbook](#) by Jake VanderPlas; the content is available [on GitHub](#).

The text is released under the [CC-BY-NC-ND license](#), and code is released under the [MIT license](#). If you find this content useful, please consider supporting the work by [buying the book](#)!

< [Application: A Face Detection Pipeline](#) | [Contents](#) | [Appendix: Figure Code](#) >

1 Further Machine Learning Resources

This chapter has been a quick tour of machine learning in Python, primarily using the tools within the Scikit-Learn library. As long as the chapter is, it is still too short to cover many interesting and important algorithms, approaches, and discussions. Here I want to suggest some resources to learn more about machine learning for those who are interested.

1.1 Machine Learning in Python

To learn more about machine learning in Python, I'd suggest some of the following resources:

- [The Scikit-Learn website](#): The Scikit-Learn website has an impressive breadth of documentation and examples covering some of the models discussed here, and much, much more. If you want a brief survey of the most important and often-used machine learning algorithms, this website is a good place to start.
- *SciPy, PyCon, and PyData tutorial videos*: Scikit-Learn and other machine learning topics are perennial favorites in the tutorial tracks of many Python-focused conference series, in particular the PyCon, SciPy, and PyData conferences. You can find the most recent ones via a simple web search.
- [Introduction to Machine Learning with Python](#): Written by Andreas C. Mueller and Sarah Guido, this book includes a fuller treatment of the topics in this chapter. If you're interested in reviewing the fundamentals of Machine Learning and pushing the Scikit-Learn toolkit to its limits, this is a great resource, written by one of the most prolific developers on the Scikit-Learn team.
- [Python Machine Learning](#): Sebastian Raschka's book focuses less on Scikit-learn itself, and more on the breadth of machine learning tools available in Python. In particular, there is some very useful discussion on how to scale Python-based machine learning approaches to large and complex datasets.

1.2 General Machine Learning

Of course, machine learning is much broader than just the Python world. There are many good resources to take your knowledge further, and here I will highlight a few that I have found useful:

- *[Machine Learning](#)*: Taught by Andrew Ng (Coursera), this is a very clearly-taught free online course which covers the basics of machine learning from an algorithmic perspective. It assumes undergraduate-level understanding of mathematics and programming, and steps through detailed considerations of some of the most important machine learning algorithms. Homework assignments, which are algorithmically graded, have you actually implement some of these models yourself.
- *[Pattern Recognition and Machine Learning](#)*: Written by Christopher Bishop, this classic technical text covers the concepts of machine learning discussed in this chapter in detail. If you plan to go further in this subject, you should have this book on your shelf.
- *[Machine Learning: a Probabilistic Perspective](#)*: Written by Kevin Murphy, this is an excellent graduate-level text that explores nearly all important machine learning algorithms from a ground-up, unified probabilistic perspective.

These resources are more technical than the material presented in this book, but to really understand the fundamentals of these methods requires a deep dive into the mathematics behind them. If you're up for the challenge and ready to bring your data science to the next level, don't hesitate to dive-in!

< [Application: A Face Detection Pipeline](#) | [Contents](#) | [Appendix: Figure Code](#) >