

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 highlight a few that I have found useful:

Machine Learning

Taught by Andrew Ng (Coursera), this is a very clearly taught, free online course covering 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!

Symbols

%automagic, 19
%cpaste, 11
%debug, 22
%history, 16
%lprun, 28
%ismagic, 13
%magic, 13
%matplotlib, 219
%memit, 29
%mode, 20-22
%mprun, 29
%paste, 11
%prun, 27
%run, 12
%time, 25-27
%timeit, 12, 25-27
& (ampersand), 77
* (asterisk), 7
: (colon), 44
? (question mark), 3
?? (double question mark), 5
_ (underscore) shortcut, 15
| (operator), 77

A

absolute value function, 54
aggregate() method, 166
aggregates
 computed directly from object, 57
 multidimensional, 60
 summarizing set of values with, 61
aggregation (NumPy), 58-63
 minimum and maximum, 59

 multidimensional aggregates, 60
 presidents average height example, 61
 summing the values in an array, 59
 various functions, 61
aggregation (Pandas), 158-170
 groupby() operation, 161-170
 MultiIndex, 140
 Planets dataset for, 159
 simple aggregation, 159-161
Akaike information criterion (AIC), 487, 489
Albers equal-area projection, 303
algorithmic efficiency
 big-O notation, 92
 dataset size and, 85
ampersand (&), 77
Anaconda, xiv
and keyword, 77
annotation of plots, 268-275
 arrows, 272-275
 holidays/US births example, 269
 transforms and text position, 270-272
APIs (see Estimator API)
append() method, Pandas vs. Python, 146
apply() method, 167
arithmetic operators, 52
arrays
 accessing single rows/columns, 45
 arithmetic operators, 52
 attributes, 42
 basics, 42
 Boolean, 73-75
 broadcasting, 63-69
 centering, 68
 computation on, 50-58