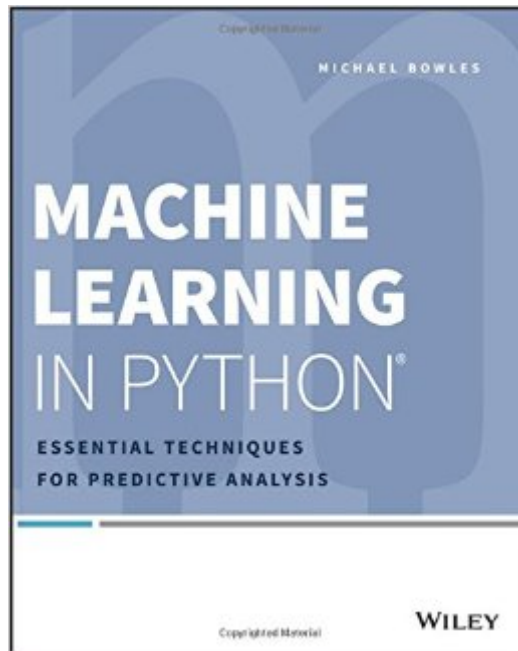


The book was found

# Machine Learning In Python: Essential Techniques For Predictive Analysis



## Synopsis

Learn a simpler and more effective way to analyze data and predict outcomes with Python Machine Learning in Python shows you how to successfully analyze data using only two core machine learning algorithms, and how to apply them using Python. By focusing on two algorithm families that effectively predict outcomes, this book is able to provide full descriptions of the mechanisms at work, and the examples that illustrate the machinery with specific, hackable code. The algorithms are explained in simple terms with no complex math and applied using Python, with guidance on algorithm selection, data preparation, and using the trained models in practice. You will learn a core set of Python programming techniques, various methods of building predictive models, and how to measure the performance of each model to ensure that the right one is used. The chapters on penalized linear regression and ensemble methods dive deep into each of the algorithms, and you can use the sample code in the book to develop your own data analysis solutions. Machine learning algorithms are at the core of data analytics and visualization. In the past, these methods required a deep background in math and statistics, often in combination with the specialized R programming language. This book demonstrates how machine learning can be implemented using the more widely used and accessible Python programming language. Predict outcomes using linear and ensemble algorithm families Build predictive models that solve a range of simple and complex problems Apply core machine learning algorithms using Python Use sample code directly to build custom solutions Machine learning doesn't have to be complex and highly specialized. Python makes this technology more accessible to a much wider audience, using methods that are simpler, effective, and well tested. Machine Learning in Python shows you how to do this, without requiring an extensive background in math or statistics.

## Book Information

Paperback: 360 pages

Publisher: Wiley; 1 edition (April 27, 2015)

Language: English

ISBN-10: 1118961749

ISBN-13: 978-1118961742

Product Dimensions: 7.4 x 0.6 x 9.2 inches

Shipping Weight: 1.2 pounds (View shipping rates and policies)

Average Customer Review: 3.9 out of 5 stars [See all reviews](#) (24 customer reviews)

Best Sellers Rank: #143,763 in Books (See Top 100 in Books) #23 in [Books > Computers &](#)

Technology > Computer Science > AI & Machine Learning > Machine Theory #31 in Books > Computers & Technology > Databases & Big Data > Data Warehousing #51 in Books > Textbooks > Computer Science > Artificial Intelligence

## Customer Reviews

I have no idea why the author didn't implement pandas or numpy instead of devoting pages to reading csv files with base python. He chose to store data in lists of lists. Why??? He chose to normalize and store data "the ol' fashioned way" instead of using numpy arrays. This makes the code less readable. (Seriously, every example begins with several lines of code simply to read the data, loop through each line, split on a ",", append to a list, etc. Using pandas' read\_csv or numpy's genfromtxt code have accomplished the same thing in one line!) The author does say, however, that he included a lot of code to help explain how the algorithms work. This is actually helpful--when it relates specifically to a machine learning algorithm. He subsequently makes use of scikit-learn objects in the examples, as he should. Why he didn't follow the same approach when reading CSV files, I have no idea. My recommendation is as follows. If you're looking for code to implement, I would not use this book for proper examples. If you're looking for an explanation of penalized regression and ensemble algorithms, this book is pretty good for that purpose in my opinion.

I'm somewhat new to the field, working on a new venture involving predictive analytics. Of the resources I'm using, I keep coming back to this book - the background, problem setup, explanation of tradeoffs, and code examples have been really good. I'm no Python developer, but I was able to narrow down to an algorithm and get it running with Anaconda and Scikit-Learn without much trouble. Compared to a lot of other resources I've seen, this is a great book for a complete view of the recent algorithms along with a how-to on getting started.

Before reading this, I think you've got to have some fundamental knowledge of statistics, data science, and Python. It's a great practical book for supervised learning applications in that it narrows things down into actionable information; you can easily insert your datasets into the example algorithms and get results. A good book on data science fundamentals is "Data Science for Business," Provost & Fawcett

Not terribly impressed. The good: I applaud the choice to limit the discussion to random forests and penalized regression instead of trying to treat every algorithm under the sun. Also, the book

certainly has a solid practical bent, while not losing track of some of the basic mathematics and concerns like overfitting and the bias-variance tradeoff. On the other hand, the book is very basic â “ little you can't get from the scikit-learn documentation â “ and some of the recommendations are bizarre (weird chartjunk instead of scatterplot grids and pairwise correlation matrices to study predictors), borderline wrong (QQ-plots to study outliers, seriously?) or incomplete (no clear distinction is made between variable importance scores and regression parameters).

The book is mostly verbatim code that repeats itself. Adding a Github repo would have helped a lot and shifted the focus to analysis instead of incremental implementation

Got this for a colleague who is implementing analytics in our product. The book addresses the topic clearly, is well structured in terms of order and flow, and covers the topics in the exact level of detail that is promised in the description. It's a great introduction with enough depth to get you started.

Code provided are horrible. Instead of using Panda, numpy and scikit author has written full length code which are boring and confusing For example :1) reading file and normalization require 3-4 line of code but author has written full page of code using loops everywhere which is boring and ineffecient 2) computing mean of a dataframe (df), can be done directly using df.mean(), but author has used for i in range(ncols): mean = summary.iloc[1, i] sd = summary.iloc[2, i]

The machine learning concepts are good and well articulated (written), but the code examples in their entirety need to be rewritten professionally. It just needs to be thrown out and rewritten from scratch. I'm very surprised at the poor code provided. Here's an example of how bad it gets: (taken from "ch02/rVMSummaryStats.py") ... type = [0]\*3 # [0, 0, 0] See the problem? 'type' is a Python keyword, and it is overwritten by this statement. Next, use of for/loops and Lists instead of numpy (i.e. instead of a vectorized approach) is inefficient and makes the reader develop bad habits. The Second Edition -- with 100% code replacement-- should be free to all who purchased this first version.

[Download to continue reading...](#)

Python: Python Programming Course: Learn the Crash Course to Learning the Basics of Python (Python Programming, Python Programming Course, Python Beginners Course) Deep Learning: Recurrent Neural Networks in Python: LSTM, GRU, and more RNN machine learning architectures in Python and Theano (Machine Learning in Python) Unsupervised Deep Learning in Python:

Master Data Science and Machine Learning with Modern Neural Networks written in Python and Theano (Machine Learning in Python) Deep Learning in Python Prerequisites: Master Data Science and Machine Learning with Linear Regression and Logistic Regression in Python (Machine Learning in Python) Convolutional Neural Networks in Python: Master Data Science and Machine Learning with Modern Deep Learning in Python, Theano, and TensorFlow (Machine Learning in Python) Deep Learning in Python: Master Data Science and Machine Learning with Modern Neural Networks written in Python, Theano, and TensorFlow (Machine Learning in Python) Machine Learning in Python: Essential Techniques for Predictive Analysis Unsupervised Machine Learning in Python: Master Data Science and Machine Learning with Cluster Analysis, Gaussian Mixture Models, and Principal Components Analysis Python: Python Programming For Beginners - The Comprehensive Guide To Python Programming: Computer Programming, Computer Language, Computer Science (Machine Language) Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) A collection of Advanced Data Science and Machine Learning Interview Questions Solved in Python and Spark (II): Hands-on Big Data and Machine ... Programming Interview Questions) (Volume 7) Beginning Python Programming: Learn Python Programming in 7 Days: Treading on Python, Book 1 Python: Python Programming For Beginners - The Comprehensive Guide To Python Programming: Computer Programming, Computer Language, Computer Science Learn Python in One Day and Learn It Well: Python for Beginners with Hands-on Project. The only book you need to start coding in Python immediately Maya Python for Games and Film: A Complete Reference for Maya Python and the Maya Python API Deep Learning: Natural Language Processing in Python with GLoVe: From Word2Vec to GLoVe in Python and Theano (Deep Learning and Natural Language Processing) Deep Learning: Natural Language Processing in Python with Word2Vec: Word2Vec and Word Embeddings in Python and Theano (Deep Learning and Natural Language Processing Book 1) Predictive Analytics with Microsoft Azure Machine Learning 2nd Edition Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies (MIT Press) Machine Learning with R - Second Edition - Deliver Data Insights with R and Predictive Analytics

[Dmca](#)