The book was found

Machine Learning For Hackers

Case Studies and Algorithms to Get You Started

Machine Learning for Hackers

O'REILLY®

Drew Conway &
John Myles White

PDF Adobe

DOWNLOAD EBOOK
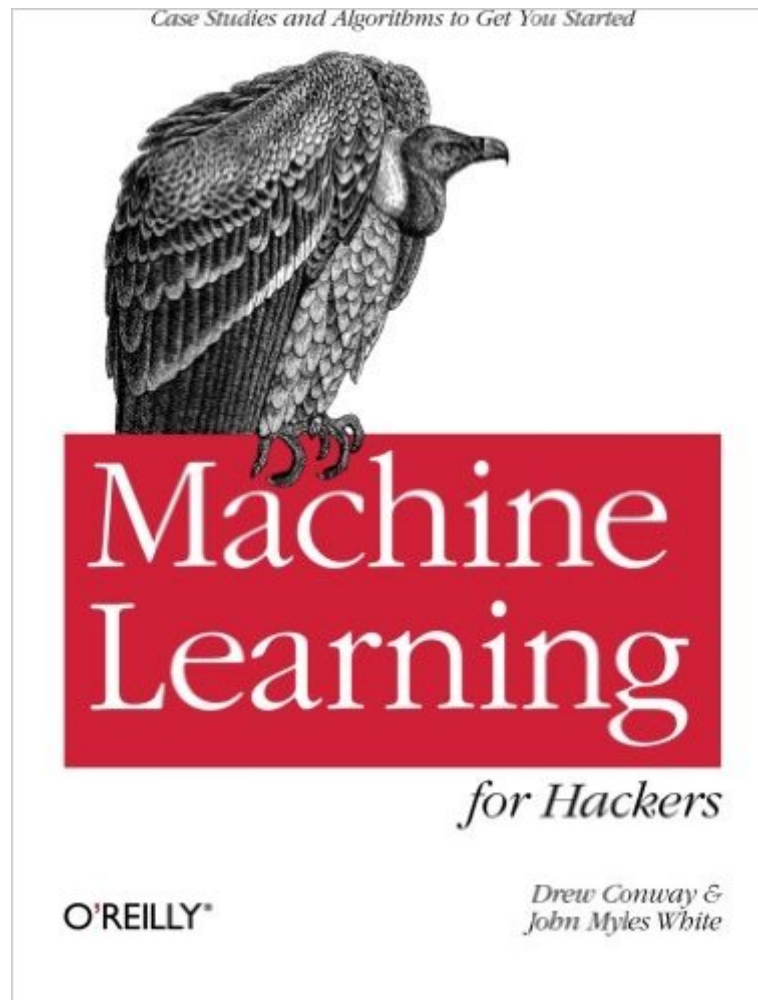
Synopsis

If youâ ™re an experienced programmer interested in crunching data, this book will get you started with machine learningâ ”a toolkit of algorithms that enables computers to train themselves to automate useful tasks. Authors Drew Conway and John Myles White help you understand machine learning and statistics tools through a series of hands-on case studies, instead of a traditional math-heavy presentation.Each chapter focuses on a specific problem in machine learning, such as classification, prediction, optimization, and recommendation. Using the R programming language, youâ ™ll learn how to analyze sample datasets and write simple machine learning algorithms. Machine Learning for Hackers is ideal for programmers from any background, including business, government, and academic research.Develop a naÃ ve Bayesian classifier to determine if an email is spam, based only on its textUse linear regression to predict the number of page views for the top 1,000 websitesLearn optimization techniques by attempting to break a simple letter cipherCompare and contrast U.S. Senators statistically, based on their voting recordsBuild a â œwhom to followâ • recommendation system from Twitter data

Book Information

Paperback: 324 pages

Publisher: O'Reilly; 1st edition (February 25, 2012)

Language: English

ISBN-10: 1449303714

ISBN-13: 978-1449303716

Product Dimensions: 7 x 0.8 x 9.2 inches

Shipping Weight: 1.4 pounds (View shipping rates and policies)

Average Customer Review: 3.2 out of 5 starsÂ Â See all reviewsÂ (26 customer reviews)

Best Sellers Rank: #307,541 in Books (See Top 100 in Books)   #31 inÂ Books > Computers & Technology > Computer Science > AI & Machine Learning > Natural Language Processing   #48 inÂ Books > Computers & Technology > Computer Science > AI & Machine Learning > Machine Theory   #75 inÂ Books > Textbooks > Computer Science > Algorithms

Customer Reviews

I used this book to teach students about data mining and machine learning with a hands-on approach. I intended it to be used as a book for the students to rely and fall back on. It is not suited well for that purpose.Pros: The book is affordable and nicely written. The authors take great care in making the book useful and entertaining and one can immediately start putting things into practise.

Also, the R examples are interesting and by itself motivating.Cons: The book has a couple of very grievous errors, that make me wonder the authors understand the subject matter. This is especially striking in the chapters on PCA and Multidimensional Scaling (which I covered in some depth in the class), but also to a lesser degree in other parts of the book that I have read more thoroughly (like optimization and linear and nonlinear regression). Many errors are not typos or simple mistakes but seem to be proof of a profound misunderstanding of concepts by the authors. I am sorry to be so blunt, but one should not write a book about topics that one is not intimate with. Given that the book is probably quite successful, it propagates error into a community whose members may not have the statistical background to spot the errors immediately. Some methods used in the book are quite hard to understand even for graduate students and to be so nonchalant about the underlying theory can be dangerous. I realize that the book is intended to be superficial with regards to mathematical or conceptual depth, but this combined with some of the presented high-level techniques can easily backfire when people are given the tools, but not the understanding.

In Machine Learning for Hackers by Drew Conway and John Myles White, the reader is introduced to a number of techniques useful for creating systems that can understand and make use of data. While the book has solid topical material and is written in a fluid and easy to read manner, I don't feel that this book is really for hackers, unless the definition of hacker is vastly different from "programmer".Much of the text is taken up explaining how to parse strings, change dates, and otherwise munge data into shape to be operated on by statistical functions provided by R. In fact, there is so much of the book in that fashion that I end up skipping through large portions to get back to something that is worth spending time reading about. I can't understand why a programmer would need significant education in string parsing. I was also put off by the vast amount of text explaining basic statistics. Maybe a recent computer science graduate is simply the wrong reader for this book?I think it is certainly possible to learn the basic principles of machine hacking from this book, and even to put them to good use with R in the same manner displayed in the examples. Indeed, the code and data available for this book would be very useful as prep for an introductory course at an academic institution. To make the best use of the text, you really should be sitting at your computer, reading the text side by side with the code, and operating on the data with R as instructed to do.Personally, I found that wading through this text wasn't enjoyable it due to the lack of density of material at the depth I was looking for. Other readers may find it is just right for them, but I suspect those readers would not be hackers, contrary to the implication of the title.

I started my journey in the machine learning / data mining field thanks to curiosity generated by Toby Segaran's classic Programming Collective Intelligence: Building Smart Web 2.0 Applications. The book by Drew Conway and John White continues in the same excellent tradition. It presents case studies which are interesting enough that you can appreciate them without too much domain knowledge and without getting too deep into technical nitty-gritty. At the same time, the case studies are meaty enough that you can adapt them to real life problems and hack together a quick working prototype in your practice.By many estimates (and my own experience), 80% of time in machine learning is spent in data cleaning and exploratory data analysis. This book has very good coverage of both areas. Authors use Hadley Wickham's excellent packages viz. ggplot2, plyr and reshape2. If you are doing serious exploratory data analysis in R, these packages are a must and the book does a great job in showing them in action.The reason I suffixed the review with 'if you know a little R' is that data cleansing requires one to be fairly comfortable with somewhat arcane R syntax. If you don't know any R at all, it would be helpful to get a more gentle introduction - such as R Cookbook (O'Reilly Cookbooks) - before you tackle this book.In summary, this is an excellent 2nd book on R to have as you try to deploy machine learning in real life.

Download to continue reading...

Machine Learning for Hackers 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: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) Unsupervised Machine Learning in Python: Master Data Science and Machine Learning with Cluster Analysis, Gaussian Mixture Models, and Principal Components Analysis Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms Foundations of Machine Learning (Adaptive Computation and Machine Learning series) Introduction to Machine Learning (Adaptive Computation and Machine Learning series) Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning series) Bioinformatics: The Machine Learning Approach, Second Edition (Adaptive Computation

and Machine Learning) First-Time Machine Applique: Learning to Machine Applique in Nine Easy Lessons 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) The City of Tomorrow: Sensors, Networks, Hackers, and the Future of Urban Life (The Future Series) Hackers: Heroes of the Computer Revolution: 25th Anniversary Edition The Innovators: How a Group of Hackers, Geniuses, and Geeks Created the Digital Revolution How to Stop E-Mail Spam, Spyware, Malware, Computer Viruses and Hackers from Ruining Your Computer or Network: The Complete Guide for Your Home and Work Thinking Security: Stopping Next Year's Hackers (Addison-Wesley Professional Computing Series)