

f (<https://www.facebook.com/AnalyticsVidhya>)

t (<https://twitter.com/analyticsvidhya>)

g+ (<https://plus.google.com/+Analyticsvidhya/posts>)

in (<https://www.linkedin.com/groups/Analytics-Vidhya-Learn-everything-about-5057165>)



(<https://www.analyticsvidhya.com>)



(<https://datahack.analyticsvidhya.com/contest/the-strategic-monk/>)

Home (<https://www.analyticsvidhya.com/>) > Machine Learning (<https://www.analyticsvidhya.com/blog/category/machine-lear...>)

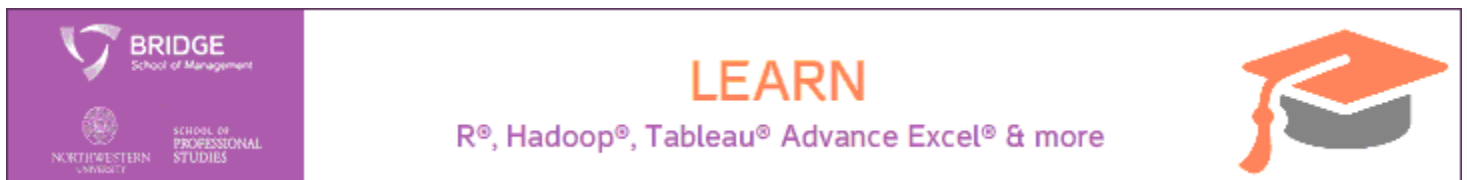
18 New Must Read Books for Data Scientists on R and Python

MACHINE LEARNING (<https://www.analyticsvidhya.com/blog/category/machine-learning/>) PANDAS

(<https://www.analyticsvidhya.com/blog/category/python-2/pandas/>) PYTHON

(<https://www.analyticsvidhya.com/blog/category/python-2/>) R (<https://www.analyticsvidhya.com/blog/category/r/>)

n/sharer.php?u=<https://www.analyticsvidhya.com/blog/2016/10/18-new-must-read-books-for-data-scientists-on-r-and-Python>) **t** (<https://twitter.com/home?status=18%20New%20Must%20Read%20Books%20for%20Data%20Scientists%20on%20R%20and%20Python+https://www.analyticsvidhya.com/blog/2016/10/18-new-must-read-books-for-data-scientists-on-r-and-python/>) **g+** (<https://plus.google.com/share?url=https://www.analyticsvidhya.com/blog/2016/10/18-new-must-read-books-for-data-scientists-on-r-and-python/>) **p:** (<https://pinterest.com/pin/create/button/?url=https://www.analyticsvidhya.com/blog/2016/10/18-new-must-read-books-for-data-scientists-on-r-and-python/&description=18%20New%20Must%20Read%20Books%20for%20Data%20Scientists%20on%20R%20and%20Python>)



([http://admissions.bridgesom.com/pba-new/?](http://admissions.bridgesom.com/pba-new/?utm_source=AV&utm_medium=BannerInline&utm_campaign=AVBanner20August)

[utm_source=AV&utm_medium=BannerInline&utm_campaign=AVBanner20August](http://admissions.bridgesom.com/pba-new/?utm_source=AV&utm_medium=BannerInline&utm_campaign=AVBanner20August))

Introduction

"It's called reading. It's how people install new software into their brain"

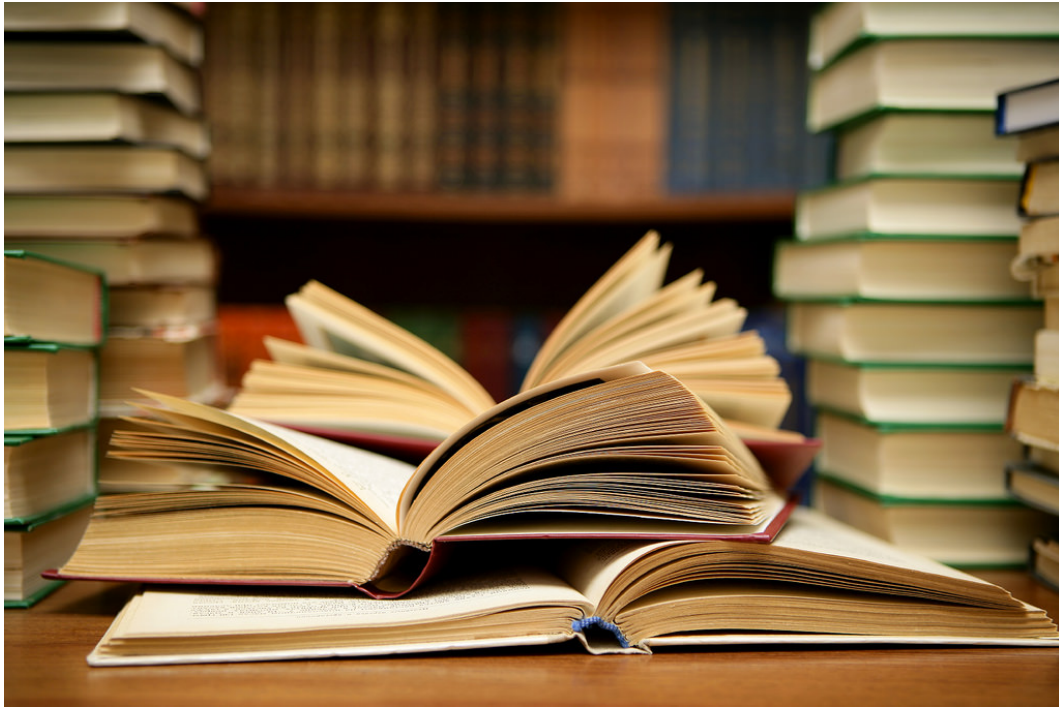
Personally, I haven't learnt as much from videos & online tutorials as much I've learnt from books. Until this very moment, my tiny wooden shelf has enough books to keep me busy this winter.

Understanding machine learning & data science is easy. There are numerous open courses which you can take up right now and get started. But, acquiring in-depth knowledge of a subject requires extra effort. For example: You might quickly understand how does a random forest work, but understanding the logic behind it's working would require extra efforts.

The confidence of questioning the logic comes from reading books. Some people easily accept the status quo. On the other hand, some curious ones challenge & say, "Why can't it be done the other way?" That's where such people discover new ways of executing a task. Almost, every data scientist I've come across in person, on AMAs, on published interviews, each one of them have emphasized the inevitable role of books in their lives.

Here is a list of books on doing machine learning / data science in R and Python which I've come across in last one year. Since reading is a good habit, with this post, I want pass this habit to you. For each book, I've written a summary to help you judge its relevance. Happy reading!

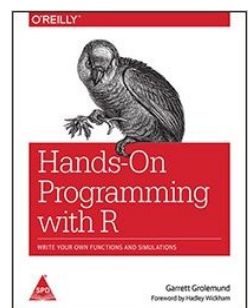
Disclosure: *The amazon links in this article are affiliate links. If you buy a book through this link, we would get paid through Amazon. This is one of the ways for us to cover our costs while we continue to create these awesome articles. Further, the list reflects our recommendation based on content of book and is no way influenced by the commission.*



R for Data Science

Hands-on Programming with R

This book is written by Garrett Grolmund. It is best suited for people new to R. Learning to write functions & loops empowers you to do much more in R, than just juggling with packages. People think, R packages can let them avoid writing functions & loops, but it isn't a sustainable approach. This book introduces you to details of R programming environment using interesting projects like weighted dice, playing cards, slot machine etc. The book language is simple to understand and examples can be reproduced easily.



Available: Buy Now (https://www.amazon.in/gp/product/9351107280/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=9351107280&linkCode=as2&tag=analyvidhy-21&linkId=1165c08ce5a0e20fee8951150fbe25f2)

R for Everyone: Advanced Analytics and Graphics

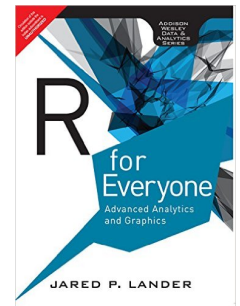
This book is written by Jared P. Lander. It's a decent book covering all aspects of data science such as data visualization, data manipulation, predictive modeling, but not in as much depth. You can understand as, it covers a wide breath of topic and misses out on details of each. Precisely, it

emphasizes on the usage criteria of algorithms and one example each showing its implementation in R. This books should be brought by people who are more inclined towards understand practical side of algorithms.

Available: Buy Now

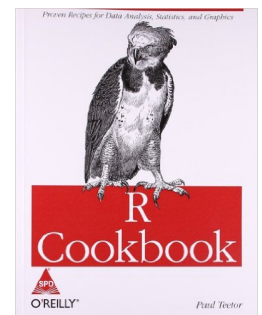
([https://www.amazon.in/gp/product/B00L6KCN1M/ref=as_li_tl?](https://www.amazon.in/gp/product/B00L6KCN1M/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=B00L6KCN1M&linkCode=as2&tag=analyvidhy-21&linkId=37cbdb89f1aba830984d88a003b3c1fc)

[ie=UTF8&camp=3638&creative=24630&creativeASIN=B00L6KCN1M&linkCode=as2&tag=analyvidhy-21&linkId=37cbdb89f1aba830984d88a003b3c1fc](https://www.amazon.in/gp/product/B00L6KCN1M/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=B00L6KCN1M&linkCode=as2&tag=analyvidhy-21&linkId=37cbdb89f1aba830984d88a003b3c1fc))



R Cookbook

This book is written by Teetor Paul. It comprises of several tips, recipes to help people overcome daily struggles in data pre-processing and manipulation. Many a times, we are stuck in a situation where we know very well, what needs to be done. But, how it needs to be done becomes a mammoth challenge. This books solves the problem. It doesn't have theoretical explanation of concepts, but focuses on how to use them in R. It covers a wide range of topics such as probability, statistics, time series analysis, data pre-processing etc.

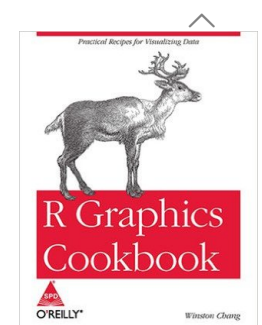


Available: Buy Now ([https://www.amazon.in/gp/product/9350233797/ref=as_li_tl?](https://www.amazon.in/gp/product/9350233797/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=9350233797&linkCode=as2&tag=analyvidhy-21&linkId=43e29866358b612d1703f9e83a1bb767)

[ie=UTF8&camp=3638&creative=24630&creativeASIN=9350233797&linkCode=as2&tag=analyvidhy-21&linkId=43e29866358b612d1703f9e83a1bb767](https://www.amazon.in/gp/product/9350233797/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=9350233797&linkCode=as2&tag=analyvidhy-21&linkId=43e29866358b612d1703f9e83a1bb767))

R Graphics Cookbook

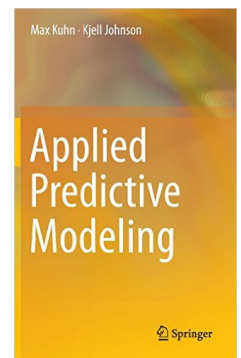
This book is written by Winston Chang. Data visualization enables a person to express & analyze their findings using shapes & colors, not just in tables. Having a solid understanding of charts, when to use which chart, how to customize a chart and make it look good, is a key skill of a data scientist. This book doesn't bore you with theoretical knowledge, but focuses on building them in R using sample data sets. It focuses on ggplot2 package to undertake all visualization activities.



Available: Buy Now (https://www.amazon.in/gp/product/9350239825/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=9350239825&linkCode=as2&tag=analyvidhy-21&linkId=e68740809122e0fb1abb977b72cd0361)

Applied Predictive Modeling

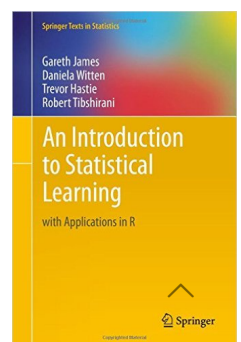
This book is written by Max Kuhn and Kjell Johnson. Max Kuhn is none other than creator of caret package too. It's one of the best book comprising a blend of theoretical and practical knowledge. It discusses several crucial machine learning topics such as over-fitting, feature selection, linear & non-linear models, trees methods etc. Needless to say, it demonstrates all these algorithms using caret package. Caret is one of the powerful ML package contributed in CRAN library.



Available: Buy Now (https://www.amazon.in/gp/product/B00K15TZU0/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=B00K15TZU0&linkCode=as2&tag=analyvidhy-21&linkId=8ebceec36330ac208779d6d456d49923)

Introduction to Statistical Learning

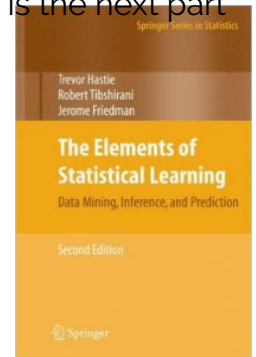
This book is written by a team of authors including Trevor Hastie and Robert Tibshirani. It is one of the most detailed book on statistical modeling. Also, it's available for free. It comprises of in-depth explanation of topics such as linear regression, logistic regression, trees, SVM, unsupervised learning etc. Since it's the introduction, the explanations are quite easy and any newbie can easily follow it. Thus, I recommended this book to all people who are new to machine learning in R. In addition, several practice exercises in this book just adds cherry on top.



Available: Buy Now (https://www.amazon.in/gp/product/B01IBM7790/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=B01IBM7790&linkCode=as2&tag=analyvidhy-21&linkId=g8f164852072597e10af54a4c1d606f1)

Elements of Statistical Learning

This book is written by Trevor Hastie, Robert Tibshirani and Jerome Friedman. This is the next part of 'Introduction to Statistical Learning'. It comprises of more advanced topics, therefore I would suggest you not to directly jump to it. This book is best suited for people familiar with basics of machine learning. It talks about shrinkage methods, different linear methods for regression, classification, kernel smoothing, model selection etc. It's a must read book for people who want to understand ML in depth.



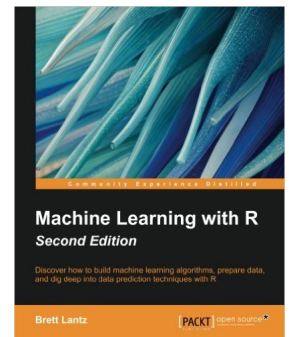
Available: Buy Now

([https://www.amazon.in/gp/product/0387848576/ref=as_li_tl?](https://www.amazon.in/gp/product/0387848576/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=0387848576&linkCode=as2&tag=analyvidhy-21&linkId=d46764655a1462258ba14ae97cb2c92d)

[ie=UTF8&camp=3638&creative=24630&creativeASIN=0387848576&linkCode=as2&tag=analyvidhy-21&linkId=d46764655a1462258ba14ae97cb2c92d](https://www.amazon.in/gp/product/0387848576/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=0387848576&linkCode=as2&tag=analyvidhy-21&linkId=d46764655a1462258ba14ae97cb2c92d))

Machine Learning with R

This book is written by Brett Lantz. I am impressed by the simplicity of this author's way of explaining concepts. It's a book on machine learning which is easy to understand, and would provide you a lot of knowledge about their practical aspects too. Algorithms such as Bagging, Boosting, SVM, Neural Network, Clustering etc are discussed by solving respective case studies. These case studies will help you understand the real world usage of these algorithms. In addition, knowledge of ML parameters is also discussed.



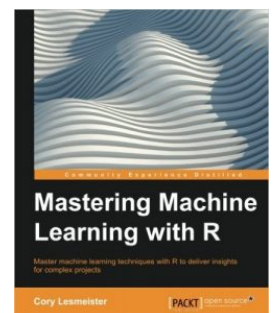
Available: Buy Now ([https://www.amazon.in/gp/product/1784393908/ref=as_li_tl?](https://www.amazon.in/gp/product/1784393908/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1784393908&linkCode=as2&tag=analyvidhy-21&linkId=e7f60f51c3af4e56b2c14f75948f0779)

[ie=UTF8&camp=3638&creative=24630&creativeASIN=1784393908&linkCode=as2&tag=analyvidhy-21&linkId=e7f60f51c3af4e56b2c14f75948f0779](https://www.amazon.in/gp/product/1784393908/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1784393908&linkCode=as2&tag=analyvidhy-21&linkId=e7f60f51c3af4e56b2c14f75948f0779))



Mastering Machine Learning with R

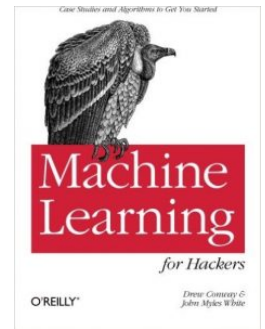
This book is written by Cory Lesmeister. It is best suited for everyone who want to master R for machine learning purposes. It comprises of all (almost) algorithms and their execution in R. Alongside, this book will introduce you to several R packages used for ML including the recently launched H2o package. It's a book which features latest advancements in ML forte, hence I'd suggest it to be read by every R user. However, you can't expect to learn advanced ML techniques like Stacking from this book.



Available: Buy Now (https://www.amazon.in/gp/product/178398452X/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=178398452X&linkCode=as2&tag=analyvidhy-21&linkId=a2216ff9e6522203a8996aa7f667f3f8)

Machine Learning for Hackers

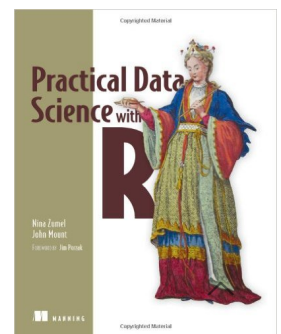
This book is written by Drew Conway and John Myles White. It's a relatively shorter book than others, but aptly brings out sheer importance of every topic discussed. After reading this book, I realized that the author's mindset is not to go deep in a topic, still making sure to cover important details. For enhanced understanding, the author also demonstrates several used cases, while solving which, explains the underlying methods too. It's a good read for everyone who'd like to learn something new about ML.



Available: Buy Now (https://www.amazon.in/gp/product/9350236745/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=9350236745&linkCode=as2&tag=analyvidhy-21&linkId=6c4791cd43a2f04e704e0decc5b12eae)

Practical Data Science with R

This book is written by Nina Zumel & John Mount. As the name suggests, this book focuses on using data science methods in real world. It's different in itself. None of the books listed above, talks about real world challenges in model building, model deployment, but it does. The author doesn't move her focus from establishing a connect between theoretical world of ML and its impact on real world activities. It's a must read for freshers who are yet to enter analytics industry.

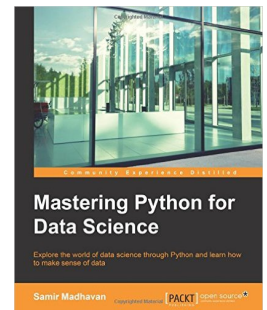


Available: Buy Now (https://www.amazon.in/gp/product/935119437X/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=935119437X&linkCode=as2&tag=analyvidhy-21&linkId=fbada304072010fe967fe412e4cf314d)

Python for Data Science

Mastering Python for Data Science

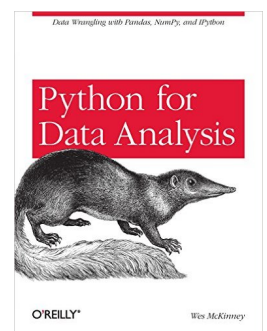
This book is written by Samir Madhavan. This book starts with an introduction to data structures in Numpy & Pandas and provides a useful description of importing data from various sources into these structures. You will learn to perform linear algebra in Python and make analysis by using inferential statistics. Later, the book takes onto the advanced concepts like building a recommendation engine, high-end visualization using Python, ensemble modeling etc.



Available: Buy Now (https://www.amazon.in/gp/product/1784390151/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1784390151&linkCode=as2&tag=analyticsvidhy-21&linkId=5220f3e71dcf44b2985a3f8578deb9f2)

Python for Data Analysis

Want to get started with data analysis with Python? Get your hands on this data analysis guide by W Mckinney, the main author of Pandas library. There isn't any online course as comprehensive as this book. This book covers all aspects of data analysis from manipulating, processing, cleaning, visualization and crunching data in Python. If you are a new to data science python, it's a must read for you. It's power-packed with case studies from various domains.



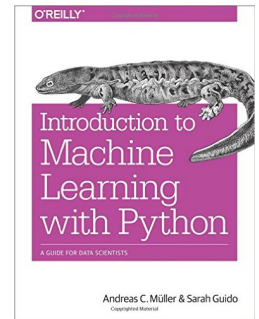
Available: Buy Now (https://www.amazon.in/gp/product/B009NLMB8Q/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=B009NLMB8Q&linkCode=as2&tag=analyticsvidhy-21&linkId=457cbe18457251e71d110a26fd444956)



Introduction to Machine Learning with Python

This book is written by Andreas Muller and Sarah Guido. It's meant to help beginners to get started with machine learning. It teaches to build ML models in python scikit-learn from scratch. It assumes no prior knowledge, hence it's best suited for people with no prior python or ML knowledge. In addition, it also covers advanced methods for model evaluation and parameter tuning, methods for working with text-data, text -specific processing techniques etc.

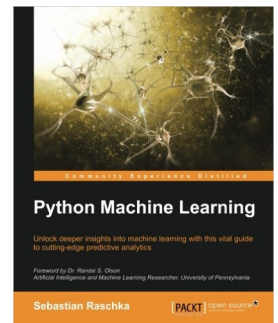
Available: Buy Now (https://www.amazon.in/gp/product/1449369413/ref=as_li_tl)



ie=UTF8&camp=3638&creative=24630&creativeASIN=1449369413&linkCode=as2&tag=analyvidhy-21&linkId=03e8851e7f1ab07708e68e27877f8b3e)

Python Machine Learning

This book is written by Sebastian Raschka. It's one of the most comprehensive book's I've found on ML in Python. The author explains every crucial detail we need to know about machine learning. He takes a stepwise approach in explaining the concepts supported by various examples. This book cover topics such as neural networks, clustering, regression, classification, ensemble etc. It's a must read book for everyone keen to master ML in python.

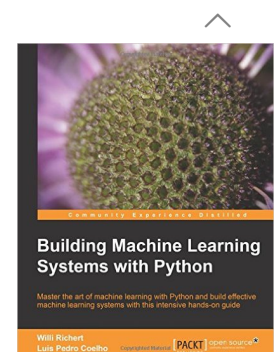


Available: Buy Now ([https://www.amazon.in/gp/product/1783555130/ref=as_li_tl?](https://www.amazon.in/gp/product/1783555130/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1783555130&linkCode=as2&tag=analyvidhy-21&linkId=aa1b34f7ced9caa8503b226ef216281f)

ie=UTF8&camp=3638&creative=24630&creativeASIN=1783555130&linkCode=as2&tag=analyvidhy-21&linkId=aa1b34f7ced9caa8503b226ef216281f)

Building Machine Learning Systems with Python

This book is written by Willi Richert, Luis Pedro Coelho. In this book the authors have chosen a path of, starting with basics, explaining concepts through projects and ending on a high note. Therefore, I'd suggest this book to newbie python machine learning enthusiasts. It covers topics like image processing, recommendation engine, sentiment analysis etc. It's easy to understand and fast to implement text book.



Available: Buy Now

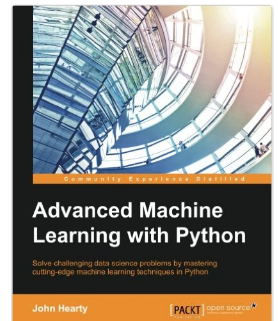
([https://www.amazon.in/gp/product/1784392774/ref=as_li_tl?](https://www.amazon.in/gp/product/1784392774/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1784392774&linkCode=as2&tag=analyvidhy-21&linkId=aa1b34f7ced9caa8503b226ef216281f)

ie=UTF8&camp=3638&creative=24630&creativeASIN=1784392774&linkCode=as2&tag=analyvidhy-21&linkId=aa1b34f7ced9caa8503b226ef216281f)

21&linkId=fb7fbc34ba94a49ada30fd7cod3eg103)

Advanced Machine Learning with Python

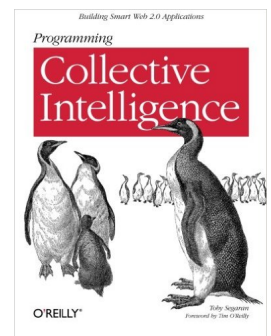
This book is written by John Hearty. It's a definite read for every machine learning enthusiasts. It lets you rise above the basics of ML techniques and dive into unsupervised methods, deep belief networks, Auto encoders, feature engineering techniques, ensembles etc. It's definitely a book you would want to read to improve your ranks in machine learning competitions. The author lays equal emphasis on theoretical as well practical aspects of machine learning.



Available: Buy Now (https://www.amazon.in/gp/product/1784398632/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=1784398632&linkCode=as2&tag=analyvidhy-21&linkId=ea8e7291ed71efb68254bedbe3a22b12)

Programming Collective Intelligence

This book is written by Toby Segaran. With an interesting title, this book is meant to introduce you to several ML algorithms such as SVM, trees, clustering, optimization etc using interesting examples and used cases. This book is best suited for people new to ML in python. Python, known for its incredible ML libraries & support should make it easy for you to learn these concepts faster. Also, the chapters include exercises for practice to help you develop better understanding.



Available: Buy Now (https://www.amazon.in/gp/product/8184043708/ref=as_li_tl?ie=UTF8&camp=3638&creative=24630&creativeASIN=8184043708&linkCode=as2&tag=analyvidhy-21&linkId=57a22fdafa62fcdag90cc12f336d227)

End Notes