

Data Science Books

Meetup #6 Data Science Campinas

Ivan Marin

Areas

- Statistical Learning/Machine Learning
- Time Series
- Pattern Recognition
- Statistics

Statistical Learning

An Introduction to Statistical Learning,

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (University of Southern California)

<http://www-bcf.usc.edu/~gareth/ISL/>

- Free!
- Introductory
- Uses R for examples
- 2015 (6th edition)
- Good for: introduction to statistical models and machine learning

Statistical Learning

Elements of Statistical Learning,

Trevor Hastie, Robert Tibshirani, Jerome Friedman (University of Southern California)

<https://web.stanford.edu/~hastie/ElemStatLearn/>

- Free!
- Medium/Advanced
- Uses R for examples
- 2009 (12th edition)
- Good for: in depth statistical basis of machine learning

Machine Learning

Machine Learning,

Kevin Murphy (Research Scientist at Google)

<https://mitpress.mit.edu/books/machine-learning-0>

- Expensive
- Probabilistic/Bayesian
- Covers a lot: Advanced
- 2012
- Good for: everything and the kitchen sink

Pattern Recognition

Pattern Recognition and Machine Learning,

Christopher Bishop (Microsoft Research)

<http://www.springer.com/us/book/9780387310732>

- Expensive (but there are ways...)
- Includes pattern recognition view in machine learning
- Medium/Advanced
- 2006
- Good for: overview and textbook, including Deep Learning

Time Series

Forecasting: principles and practice

Rob J Hyndman, George Athana-sopou-los (Monash University, Australia)

<https://www.otexts.org/fpp>

- Free!
- Includes EDA for Time series
- Introductory, uses R
- 2017
- Good for: Introduction on the subject

Statistics

Engineering Statistics,

NIST (US Department of Commerce)

<http://www.itl.nist.gov/div898/handbook/>

- Free!
- EDA, measurements, distributions, methods, Statistical models
- Introductory/Medium
- 2012
- Good for: reference manual and that things you can never remember
- (n or n-1?)

Deep Learning

CS231n: Convolutional Neural Networks for Visual Recognition

University of Stanford

<http://cs231n.stanford.edu/>

- Free!
- Not exactly a book, but lots of materials
- Introductory/Medium/High - it's a long course
- 2017
- Good for: introduction with examples in Deep Learning

Bonus: O'Reilly

O'Reilly books

- Lots and lots of areas
 - Including business
- Good books (Pandas, Spark)
- Bad books (you name it)
- Expensive (but ask your employer, you may have access!)

Bonus: O'Reilly

Cutting corners to meet arbitrary management deadlines



Essential

Copying and Pasting from Stack Overflow

O'REILLY®

The Practical Developer
@ThePracticalDev

The internet will make those bad words go away



Essential

Googling the Error Message

O RLY?

The Practical Developer
@ThePracticalDev

Software can be chaotic, but we make it work



Expert

Trying Stuff Until it Works

O RLY?

The Practical Developer
@ThePracticalDev