Data Science & Machine Learning Bibliography

January 18, 2022

Prepared by Daniel D. Gutierrez – UCLA Extension

Data Science Books

- Gutierrez, Daniel D. *Machine Learning and Data Science: An Introduction to Statistical Learning Methods with R*, Technics Publications, 2015 [required textbook for my classes at UCLA Extension, great for beginners].
- O'Neil, Schutt. *Doing Data Science*, O'Reilly, 2014 [interesting perspectives about the data science industry, some math, some R code, some Python code].

Machine Learning Books

- Hastie, Tibshirani, Friedman. *Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Edition*, Springer, 2016 [known as the "machine learning" Bible, available as free PDF download].
- James, Witten, Hastie, Tibshirani. *Introduction to Statistical Learning with Applications in R*, Springer, 2013 [excellent text after completing "intro" class, available as free PDF download].
- Conway, white. *Machine Learning for Hackers*, O'Reilly, 2012 [each chapter is a case study, excellent style R code].
- Kuhn, Johnson. *Applied Predictive Modeling*, Springer, 2013 [uses R for machine learning, written by creator of `caret` package].
- Sheather, Simon J. A Modern Approach to Regression with R, Springer, 2009.
- Abu-Mostafa, Magdon-Ismail, Lin. *Learning from Data*, self-published, 2012 [all theory, and mathematics of ML by Caltech professors, used in undergrad course].
- Hastie, Tibshirani, Wainwright. *Statistical Learning with Sparsity: The LASSO and Generalizations*, CRC Press, 2015.

- Takahashi, Inoue. *The Manga Guide to Regression Analysis*, No Starch Press, 2016 [alternative way to learn regression with Japanese comics, detailed calculations for optimal learning].

AI and Deep Learning Books

- Glassner, Andrew. *Deep Learning: A Visual Approach*, No Starch Press, 2021.
- Goodfellow, Bengio, Courville. *Deep Learning*, MIT Press, 2016 [most authoritative text on deep learning theory]
- Charniak, Eugene. *Introduction to Deep Learning*, MIT Press, 2018 [great introduction to DL with mathematics and Python code].
- Sejnowski, Terrence J. *The Deep Learning Revolution*, MIT Press, 2018 [non-technical overview and historical perspective].
- Stone, James V. *Artificial Intelligence Engines: A Tutorial Introduction to the Mathematics of Deep Learning*, Sebtel Press, 2019.

R Programming Books

- Baumer, Kaplan, Horton. *Modern Data Science with R, 2nd Edition*, CRC Press, 2021 ["modern" coding with R using the tidyverse].
- Hoffmann, John P., *Linear Regression Models – Applications in R*, CRC Press, 2022.
- Matloff, Norman. *The Art of R Programming: A Tour of Statistical Software Design*, No Starch Press, 2011 [good resource for learning R programming from a CS professor's perspective].
- Davies, Tilman M. *The Book of R: A First Course in Programming and Statistics*, No Starch Press, 2016 [great R programming text, plus Part III on Statistics and Probability, and Part IV on Statistical Testing and Modeling].
- Crawley, Michael J. *The R Book*, John Wiley & Sons, 2007 [lots of R code, lots of statistics in this 942 page tome].
- Pearson, Ronald J. *Exploratory Data Analysis Using R*, CRC Press, 2018.
- Wickham, Hadley. *ggplot2: Elegrant Graphics for Data Analysis*, Springer, 2009 [authoritative text by the R guru].

- Adler, Joseph. *R in a Nutshell: A Desktop Quick Reference, 2nd Edition*, O'Reilly, 2012.
- Lander, Jared P. *R for Everyone*, Addison Wesley, 2014
- Pearson, Ronald K. *Exploratory Data Analysis Using R*, CRC Press, 2018

Python Books

- Vaughan, Lee. *Real World Python*, No Starch Press, 2021.
- Liu, Yuxi (Hayden). *Python Machine Learning by Example*, Packt Publishing, 2020.
- Mukhiya, Suresh Kumar and Ahmed, Usman. *Hands-on Exploratory Data Analysis with Python*, Packt Publishing, 2020.
- Matthes, Frank. *Python Crash Course*, No Starch Press, 2019.
- VanderPlas, Jake. *Python Data Science Handbook*, O'Reilly, 2017 [excellent text for learning Python for data science and ML].
- Raschka, Mirjalili. *Python Machine Learning, 3rd Edition*, Packt Publishing, 2019 [great way to learn ML with Python, after you know Python].
- Saha, Amit. *Doing Math with Python*, No Starch Press, 2015.

Data Visualization

- Knaflic, Cole Nussbaumer. *Storytelling with Data: A Data Visualization Guide for Business Professionals*, John Wiley & Sons, 2015 [extensive guide for the final step in the data science process: data storytelling].
- Berinato, Scott. *Good Charts*, Harvard Business Review Press, 2016.
- Duarte, Nancy. *Data Story: Explain Data and Inspire Action through Story*, IDEAPRESS Publishing, 2019.

Mathematics Books

- Deisenroth, Marc P., Faisal, A.Aldo, Ong, Cheng Soon. *Mathematics for Machine Learning*, Cambridge University Press, 2020 [complete text for understanding the mathematics behind ML].

- Kneusel, Ronald T. *Math for Deep Learning*, No Starch Press, 2022 [Python based book on mathematical foundations of DL]
- Strang, Gilbert. *Linear Algebra and Learning from Data*, Wellesley Cambridge Press, 2019 [great way to learn the mathematics of ML by the master MIT professor].
- Strang, Gilbert. *Introduction to Linear Algebra, 5th Edition*, Wellesley Cambridge Press, 2016 [excellent for linear algebra foundations for ML].
- Takahashi, Inoue. *The Manga Guide to Linear Algebra*, No Starch Press, 2012 [an alternative way to learn linear algebra using Japanese manga (comics)].
- Apostol, Tom M. *Calculus Volume I: 2nd Edition*, John Wiley & Sons, 1967 [my favorite Calculus text, used for undergrad classes at Caltech].
- Apostol, Tom M. *Calculus Volume II: 2nd Edition*, John Wiley & Sons, 1969.
- Ayres, Mendelson. *Calculus: 6th Edition*, Schaum's Outlines, 2013 [great for doing solved problem sets].
- Snyman, Wilke. *Practical Mathematical Optimization: Basic Optimization Theory and Gradient-Based Algorithms*, Springer, 2018 [optimization is the basis of machine learning and deep learning].

Statistics Books

- Matloff, Normal. *Probability and Statistics for Data Science: Math+R+DATA*, CRC Press, 2020.
- Diez, Cetinkaya-Rundel, Barr. *OpenIntro Statistics, 4th Edition*, OpenIntro.org, 2019.
- Pruim, Randall. *Foundations and Applications of Statistics: An Introduction Using R*, American Mathematical Society, 2010 [820 page text with excellent R code].
- Ross, Sheldon M. *Introduction to Probability Models, 8th Edition*, Academic Press, 2003 [textbook for CS112: Computer Systems Modeling Fundamentals class at UCLA].
- Spiegelhalter, David. *The Art of Statistics: How to Learn from Data*, Basic Books, 2019.

- Gentle, James E. *Computational Statistics*, Springer, 2009 [wonderful 727 page text with some R code].
- Hoff, Peter D. *A First Course in Bayesian Statistical Methods*, Springer, 2010 [heavy on theory, mathematics, and R code].
- Albert, Jim. *Bayesian Computation with R*, Springer, 2009.
- Kurt, Will. *Bayesian Statistics the Fun Way*, No Starch Books, 2019.
- Rohatgi, Vijay K. *Statistical Inference*, Dover Books, 2003 [authoritative resource for mathematical statistics].
- Reinhart, Alex. *Statistics Done Wrong: The Woefully Complete Guide*, No Starch Press, 2015.

General Reading Books

- Christian, Brian. *The Alignment Problem: Machine Learning and Human Values*, W. W. Norton & Company, 2020.
- Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, 2015
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Broadway Books, 2017.