Books

Textbooks

- James et al. (2013), available online here. This will be the primary source for the course.
- Hastie, Tibshirani, and Friedman (2001), available online here. A more technical and comprehensive precursor to (James et al., 2013).

Statistics

- Pichler (2018): Lecture notes for the TU Chemnitz undergraduate statistics class, which is recommended for all MSc Data Science students without an undergraduate math degree.
- Williams (2010): A very lively and mathematicaly satisfying account of statistics and probability theory at the beginning graduate level.
- Efron and Hastie (2016): A very readable account of classical and modern statistical ideas, available online here.

Programming

- Grus (2015)
- Géron (2017), available online here.

Data Science

- Sutton and Barto (2018), available online here.
- Goodfellow, Bengio, and Courville (2016), available online here.
- Chollet (2018)
- Kelleher, Namee, and D’Arcy (2015)
- Schölkopf and Smola (2002)

Popular Science Books

- Bostrom (2014) Nick Bostrom, a Swedish philosopher at Oxford University, argues that if machine brains surpass human brains in general intelligence, then this new superintelligence could replace humans as the dominant lifeform on Earth.
- Domingos (2015) Outlines five tribes of machine learning: inductive reasoning, connectionism, evolutionary computation, Bayes' theorem and analogical modelling. The author explains these tribes to the reader by relating these to more familiar concepts of logic, connections made in the
brain, natural selection, probability and similarity judgements. Throughout the book, it is suggested that each different tribe has the potential to contribute to a unifying “master algorithm”.

- O’Neil (2016) O’Neil, a mathematician and former Wall Street quant, analyses how the use of big data and algorithms in a variety of fields, including insurance, advertising, education, and policing, can lead to decisions that harm the poor, reinforce racism, and amplify inequality.

- Stephens-Davidowitz (2017) Inspired by Google Trends, former Google data scientist Seth Stephens-Davidowitz reveals what can be inferred about human desires, beliefs and prejudices from analyzing the vast logs of anonymous Google searches. A fascinating, if sobering, account.


What is Data Science?

- Bühlmann and Stuart (2016). A concise take on the role of math and stats within the emerging discipline of data science centering on models, high dimensionality and heterogeneity.

- Donoho (2017). Based on a presentation at the John Tukey 100th Birthday Celebration held in Princeton 2015, this overview traces the origins of the discipline, highlighting the role of statistics in the genesis of data science.

- Carmichael and Marron (2018)


Chapter 4

- Bayes’ theorem:
  - Efron (2013) : On the occasion of the 250th anniversary of Bayes’ rule, eminent statistician Bradley Efron gives a very readable account of the dispute between Bayesians and frequentists delivered as the 85th Gibbs lecture at the 2012 Joint Mathematics Meeting.
  - McGrawne (2012), a popular science book on the history and real-world impact of Bayes’ theorem

- Breast cancer screening:
  - Hoffrage and Gigerenzer (1998): How medical professionals can be taught to perform the calculations required to apply Bayes’ rule.
  - Kerlikowske et al. (1996), Kerlikowske et al. (1996), A study determining the statistical parameters of mammography screening tests.

Chapter 8

- Tree-based methods
  - More details on optimal pruning of decision trees can be found in Breiman et al. (1984) (Chapter 10) and Ripley (1996) (Chapter 7).

Chapter 9

- Linear classifiers
  - A precursor of neural networks and machine learning in general was the pioneering work of the American statistician Frank Rosenblatt, whose perceptron was an analog device which could construct separating hyperplanes (Rosenblatt, 1958).
  - The solution of the optimization problem for finding the optimal separating hyperplane is discussed in Vapnik (2000) (Section 5.5).

- Kernel methods
An excellent introduction into kernel-based methods and support vector machines can be found in Schölkopf and Smola (2002).

References


James, Gareth et al. (2013). An Introduction to Statistical Learning – with Applications in R. corrected 7th printing. Springer. DOI: 10.1007/978-1-4614-7138-7 (cit. on p. 1).


