

# References

## Textbooks

- Pattern Recognition and Machine Learning, by Christopher M. Bishop
- Pattern Classification 2nd Edition, Duda, Hard and Stork
- Convex Optimization, Stephen Boyd and Lieven Vandenberghe
- Introduction to Linear Algebra by Gilbert Strang
- A first course in probability, Sheldon Ross, 8th edition

## Other references

- Linear and Logistic regression: [Andrew Ng, CS229 course notes](#).
- [Linear algebra review and reference: Zico Kolter](#)
- Lasso references
  - [Regression Shrinkage and Selection Via the Lasso](#)
  - [Least Angle Regression](#)
  - [Compressive Sensing Resources](#)
- Classification model metrics: [Precision and recall](#), [sensitivity and specificity](#), [ROC curves](#) (read full sections).
- Multivariate Gaussian Distribution
  - [A short reference on Multivariate Gaussian Distributions](#), by Leon Gu, CMU
  - [Multivariate Gaussian Distributions](#), by Chuong B. Do, Stanford University
- Neural networks and back-propagation.
  - [Roi Livni \(Princeton\) lecture notes on back-propagation](#).
  - [CS231n \(Stanford\) slides on back-propagation](#).
  - [Matt Gormley \(CMU\) slides on back-propagation](#)
  - [Slides on Back-propagation Learning by David S. Touretzky](#)
  - [Notes on back-propagation by Christopher Bryant \(feed forward NNs as well as deep learning\)](#)
- [Gaussian Discriminant Analysis \(GDA\)](#) and [Naive Bayes Classification](#) (Andrew Ng Lecture Notes)
- [Another link for Naive Bayes Classification](#)
- [A Probabilistic Model for Linear Regression](#)
- [Estimation Theory](#), slides [from Cristiano Porciani](#), Point [estimation notes by Herman Bennett](#)
- Density estimation (use Bishop). [Here is more if you wish to go deeper](#).
- Decision Trees
  - Chapter 8.1-8.4 from Pattern Classification 2nd Edition, Duda, Hard and Stork.
  - [Chapter on Decision Trees](#) by Lior Rokach and Oded Maimon
  - [Slides on Decision Trees](#) and Random Forest by David Sontag
- [Classical paper on Random Forest](#) by Leo Breiman
- Andrew Ng course (CS229) [lecture notes on Support Vector Machines](#) (SVMs)
- Chapters 5.1, 5.2, 5.3, 5.4, 5.5 from Convex Optimization text book.
- [A Tutorial on Principal Component Analysis](#) (PCA) by Jonathon Shlens
- [Slides on PCA](#) by Barnabás Póczos
- Lee, Daniel D., and H. Sebastian Seung. [Algorithms for non-negative matrix factorization](#). In Advances in neural information processing systems, pp. 556-562. 2001.

- Lee, Daniel D., and H. Sebastian Seung. [Learning the parts of objects by non-negative matrix factorization](#). Nature 401, no. 6755 (1999): 788.
- Y. Koren, R. M. Bell, and C. Volinsky. [Matrix factorization techniques for recommender systems](#). IEEE Computer, 42(8):30–37, 2009.
- [A Global Geometric Framework for Nonlinear Dimensionality Reduction](#), Joshua B. Tenenbaum, Vin de Silva, John C. Langford and Nonlinear [Dimensionality Reduction by Locally Linear Embedding](#), Sam T. Roweis and Lawrence K. Saul
- Yoav Freund and Robert E. Schapire. [A short introduction to boosting](#). Journal of Japanese Society for Artificial Intelligence, 14(5):771-780, September, 1999.
- [VC Dimension & The Fundamental Theorem](#), Lecture notes, Class of Roi Livni
- Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "[Generative adversarial nets](#)." In Advances in neural information processing systems, pp. 2672-2680. 2014.
- 18 [Impressive Applications of Generative Adversarial Networks](#) (GANs)

## Model Answers

- [Minor 1 model answers](#)