

CSCI 4155/6505 (2017): Machine Learning

Thomas P. Trappenberg
Dalhousie University

Acknowledgements

These lecture notes have been inspired by several great sources. I would specifically like to acknowledge the influence of Andrew Ng's lecture notes on earlier versions of this manuscript, as well as the teaching material from the NVIDIA Deep Learning Institute.

There are now a variety of excellent books on the theory of machine learning that I would like to recommend for further readings. This includes the by *Introduction to Machine Learning* by Ethem Alpaydin, 2nd edition, MIT Press 2010, and *Pattern Recognition and Machine Learning* by Christopher Bishop, Springer 2006. The standard book on RL is *Reinforcement Learning: An Introduction* by Richard Sutton and Andrew Barto, MIT press, 1998. The standard book for AI, *Artificial Intelligence: A Modern Approach* by Stuart Russell and Peter Norvig, 2nd edition, Prentice Hall, 2003, does also include some chapters on Machine Learning.

The most in-depth book on probabilistic machine learning is likely the book by Kevin Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012. The great book on Deep Learning is the book by Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT Press, 2016. These are two books that are invaluable for researchers in machine learning.

Several people have contributed considerably to this lecture notes. In particular I would thank Paul Hollensen, Patrick Connor, and Hossein Parvar.

Contents

1	Introduction	1
1.1	The basic idea behind supervised Machine Learning	1
1.2	Mathematical formulation of the supervised learning problem	5
1.3	Applied Learning: Training, validating and testing	8
1.4	Non-linear regression and high-dimensionality	12
2	Practical ML programming with Python	17
2.1	General scientific programming in Python	17
2.2	Further useful functions	20
2.3	Cross validation example from Intro	20
2.4	Classification with support vector machine using scikit-learn	22
2.5	Other classification methods including MLP with Tensorflow	23
2.6	Applying ML methods to specific problems	25
3	Regression and optimization	27
3.1	Linear regression and gradient descent	27
3.2	Error surface and challenges for gradient descent	29
3.3	Regularization: Ridge regression and LASSO	31
3.4	Nonlinear regression	34
3.5	Advanced optimization (learning)	36
4	Basic probability theory	40
4.1	Random numbers and their probability (density) function	41
4.2	Moments: mean, variance, etc.	42
4.3	Examples of probability (density) functions	45
4.4	Some advanced concepts	47
4.5	Density functions of multiple random variables	51