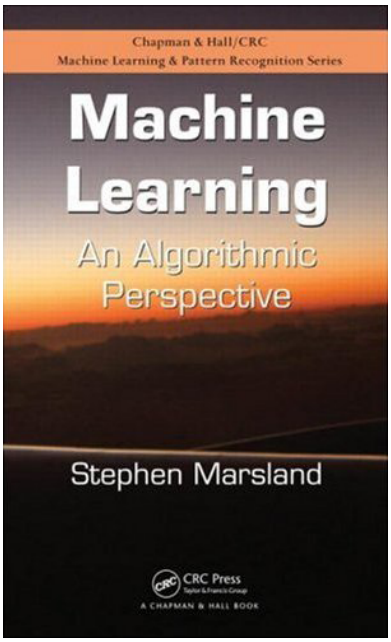
Machine Learning: An Algorithmic Perspective

STEPHEN MARSLAND



REVIEWED BY J.P. LEWIS

When several good books on a subject are available the pedagogical style of a book becomes more than a secondary consideration. This is particularly true in the case of mathematical and algorithmic subjects such as machine learning, where the level of formal rigor is a consideration. Peter Naur spent a portion of his career considering this issue. As the 'N' in the Backus-Naur formalism (BNF), one would expect Naur to champion the role of correct formal derivations in learning mathematical topics. On the contrary, in several of the chapters in his books [1], [2] Naur shows that formal approaches are at best incidental, and more often detrimental, to the learning and understanding of subjects that involve formal systems. Said differently, humans are much better at learning by example and experimentation than by attempting to follow proofs. While Naur demonstrates this in studies and observation of beginning programmers, he also illustrates the problem in a professional setting: mathematical proofs in peer-reviewed papers as short as several pages have been found to have errors after publication.

Marsland's new book *Machine Learning: An Algorithmic Perspective* takes a decisive approach to this issue, based on algorithmic experimentation. Each topic is motivated by creative examples (such as learning to dance at a nightclub) and then presented both mathematically and algorithmically. Many of the exercises require exploring and revising the code fragments in the book. There are mathematical illustrations, but no explicit proofs.

The book's example-based approach evidently effects the ordering of topics, which is occasionally odd from the perspective of someone who already has the big picture of the field. For example, the curse of dimensionality is a consideration for all machine learning approaches and thus might logically be introduced in an abstract overview chapter along with Maximum Likelihood, MAP, and so on. Instead, Marsland's book introduces it as it arises in a discussion of spline and radial basis interpolation.

The algorithmic examples in the book use the Numpy and Scipy environments in the Python language. For those not familiar, Python+Numpy is rapidly taking a place along side Matlab for the rapid prototyping of mathematical algorithms. Appealing aspects of Python are that it is a well designed and structured language with broad adoption, and the fact that it is free and open source. Experienced Matlab programmers will note many operations with similar names and behavior [3], and Numpy shares Matlab's expressiveness in representing linear algebra computations. For example, a linear discriminant example in the book is 15 lines of code, and the kernel PCA algorithm is 14 lines. On the negative side, Numpy and Scipy are still rapidly evolving and somewhat immature.

Subjects covered by *Machine Learning: An Algorithmic Perspective* include linear discriminants, neural networks, radial basis functions and splines, support vector machines, regression trees, basic probability theory and the bias-variance tradeoff, classification by neighbor neighbors, mixture models and EM, ensemble techniques, k-means, vector quantization and self organising maps, dimensionality reduction, MDS, and manifold learning, genetic algorithms, reinforcement learning, hidden Markov models, and MCMC. The book has a chapter introducing Python for Matlab and R users. A chapter on optimization initially seems out of place, though it sets the context for several other chapters.

Although several excellent and tested books on machine learning exist (e.g. [4], [5], [6]), Marsland's text stands out as the only book suited to undergraduate or Masters level teaching or equivalent self-instruction, and I expect that it will shine in this role. The book could be improved with the addition of a concluding summary chapter wherein fundamental concepts (ML, MAP, etc.) are revisited in their broadest context. After reading Naur's work [1], [2] I now believe that *all* books in the algorithmic and mathematical areas should contain strong reader advisories to watch out for typos and errors – particularly in first editions such as this.

REFERENCES

[1] P. Naur, *Computing: A Human Activity*, ACM Press 1992
[2] P. Naur, *Knowing and the Mystique of Logic and Rules*, Kluwer Academic, 1995
[3] http://www.scipy.org/NumPy_for_Matlab_Users
[4] C. Bishop, *Pattern Recogntion and Machine Learning*, Springer, 2007
[5] R. Duda, P. Hart, D. Stork, *Pattern Classification*, Wiley, New York, 2001
[6] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer Verlag, New York, 2009

ABOUT THE REVIEWER:

JOHN LEWIS: Weta Digital and Massey University. Contact him at zilla@computer.org, and see http://scribblethink.org for more information.