An introduction to Machine Learning in quantitative finance

Book review on book **"An introduction to Machine Learning in quantitative finance"**

paper authored by Ni, Dong, Zheng, Yu

Book review by C. Strange and G. dos Reis

1 Book review

The book "An Introduction to Machine Learning in Quantitative Finance" [NDZY21] by Ni, Dong, Zheng and Yu targets the currently very popular field of machine learning (ML) in finance. This topic has been gaining increasing attention in the mathematical finance research community, among quantitative finance practitioners and in graduate studies.

There are presently many books in ML but not many specific to finance. This book is not designed to be a reference book but an entry point to ML. Broadly speaking ML algorithms are commonly divided into categories (according to purpose) with the main ones being: Supervised learning, unsupervised learning and reinforcement learning (there are more). This book's main content is supervised learning tools (Ch 2-5) with two chapters on unsupervised learning (Ch 6-7) and a topical one on reinforcement learning (Ch 8). The final chapter (Ch 9) is a beginning-to-end run-through of a case-study stemming from a *Kaggle*[1] competition on a financial problem. It covers all stages of the problem from data pre-processing, feature generation, model training, tuning and selection, to the final prediction. The initial part of the book introduces ML more broadly with the financial applications appearing as the book develops. The foundational models are introduced and discussed mathematically while the more advanced ones are only briefly presented. The approach gives the reader a solid basis and broad view of available techniques without overwhelming them with mathematical rigor. Methods presented in the first five chapters are summarized by tables providing a pro/con discussion. Its is unfortunate that this systematic approach is lost past Chapter 5. Overall the proofs presented rely on calculus, statistics and probability, and not necessarily on the stochastic analysis needed for traditional mathematical finance. This means that an MSc student starting a financial program could pick up this book on their first day.

The book has an accompanying GitHub repository with commented code that can be downloaded and tested by the reader. The approach to code is not the usual 'informatics' one where one would see the multiple functions and packages used being described and discussed. The examples are mostly built using existing Python packages and the authors are not holding one's hand through the commented code. Without prior knowledge and just by inspection of the code, one would not know why the pieces come together the way they do. Take Chapter 5 as example: it is shown explicitly how neural networks work, the code is presented and every line is commented but the action of coding is not presented. The reader must meet the requirement of coding knowledge (in Python).

The leading author, Dr. Ni, is a well-known academic emerging from the research

[1] https://www.kaggle.com/

group of Prof. Terry Lyons (Oxford University) and has an excellent theoretical and applied mathematics background. Dr. Dong, Dr. Zheng and Mr. Yu are practitioners in quantitative finance from well-known financial institutions and the four form a very complementary team. The authors' motivation for the book is explained in page 2 of the book. The book collects the lecture notes of an introductory course on supervised learning delivered by Ni. Which consisted of two introductory blocks on supervised learning complemented by practice sessions and a financial case study. Further content was later added to neatly round the material, with sections on unsupervised learning as well as code produced by an algorithm engineer (Dr. Zheng).

There are a few books on ML in the finance context. For instance de Prado's book [DP18] or Dixon, Halperin & Bilokon book [DHB20]. The former is a specialized book from an author with in-depth knowledge of the financial world and problems appearing there, and ML appears as a new collection of tools that can be used to address these problems – the mantra being 'these are the issues, this is how ML can be used to tackle them'. The latter one provides, in the scope of quantitative finance and statistical ML, an in-depth discussion of supervised and reinforcement learning – unsupervised learning is left out.

In comparison to [DP18] or [DHB20], this book by Ni, Dong, Zheng and Yu can be viewed as "an introduction to ML with a tinge of finance". On the other hand, the knowledge of this book is clearly transferable and not overspecialized – one does not need deep knowledge of finance (as for [DP18] or [DHB20]) or a PhD degree to understand what is happening.

To get the most out of the book one needs a good mathematical understanding to follow some of the proofs, basic statistical knowledge and prior basic programming knowledge in Python (one is expected to have already a working Python installation with selected packages on one's machine). Although the proofs require solid knowledge in mathematics, they may be skipped on the first read.

This book's target audience is graduate students starting a finance MSc program, PhD or postdoctoral researchers that lack a view of this very specialized field, or finance practitioners wanting to learn the basics of ML without an onerous time commitment. An academic needing to build an introductory/crash course on 'ML with a tinge of finance' could very well do it using this book as a spring board or then use Chapter 9 as a basis for a an extended project.

I'm happy to have such a book available for my finance MSc and PhD students that want a quick entry point to this theme. The book is not a look-up reference but a good place to gain a basic working knowledge of ML and to prepare the reader for more advanced material.

References

[DHB20] Matthew F Dixon, Igor Halperin, and Paul Bilokon. *Machine Learning in Finance*. Springer, 2020.

[DP18] Marcos Lopez De Prado. *Advances in financial machine learning*. John Wiley & Sons, 2018.

[NDZY21] Hao Ni, Xin Dong, Jinsong Zheng, and Guangxi Yu. *An Introduction to Machine Learning in Quantitative Finance*. World Scientific, 2021.