

Commentary: Machine learning in quantitative investing – revolution or evolution?



Seth Weingram

By Seth Weingram

The topic of artificial intelligence is generating a constant stream of questions from our institutional client base. In meetings these days, it's not a matter of whether but of how quickly investors will raise the issue. That's not surprising. The prospect of AI's widespread application in all kinds of contexts has sparked hope and unease, and it's only natural that asset owners are scrambling to understand how it will affect them.

But the breathless nature of the broader societal dialogue has clouded understanding of AI's implications for investing, complicating the task of distinguishing reality from science fiction and relevance in the near term from the limitless future.

Institutional investors are wise to focus on the topic. Machine learning or ML — data-driven approaches to achieving artificial intelligence such as decision trees and neural networks — has genuinely exciting potential for quantitative managers; it significantly expands their analytical toolkit. ML can help to infer the form of a predictive relationship rather than imposing a simple assumption and to illuminate subtle structure in complex and/or “big” data.

The added analytical flexibility offers benefits across the investment process. ML can extract quantitative data inputs from qualitative information, such as in text-based sentiment analysis. In development of technical signals, ML can better exploit richness in historical returns than traditional forecasting approaches that traded off nuance for robustness. In cross-asset sentiment analysis, ML can help to illuminate subtle linkages among companies.

But markets present special challenges for ML. One is that the financial data available to train an algorithm is limited to what history provides. Another concerns the constantly changing economic and institutional environment in which the data is generated. In contrast, in training a machine to dominate a game like chess or go, we can invent more data simply by playing more games, and the rules are immutable.

Such considerations help to explain many past failures in applying ML to investing. A relative scarcity of training data and a low signal-to-noise ratio implies that blindly unleashed ML algorithms are prone to overfitting, or picking up on spurious relationships in past data that don't help in out-of-sample forecasting. Forgotten academic articles from the 1990s and long-dead web links chronicle a persistent pattern of unrealistic expectations on the part of methodologically sophisticated but market-naïve researchers.

Unfortunately, ML also makes it harder for asset owners to discriminate between informed analysis and sloppy or abusive data mining. ML's modeling flexibility increases the likelihood of picking up on spurious relationships, and its algorithmic complexity reduces visibility into what's driving its forecasts. Think of a neural network that consists of hundreds or thousands of simple functional elements. Knowledge of what each one of them does will provide little insight into the behavior of the entire system.

As a result of ML's opacity and the special challenges in its application to investing, asset owners would be wise to seek out managers who follow a set of best practices.

First, the decision whether to apply ML should be guided by deep understanding of the specific problem at hand and the relevant market context. This knowledge will inform judgment as to how much benefit there may be to ML's added flexibility relative to the risk of overfitting. Once ML has been selected for a given problem, context-specific investment knowledge will also inform decisions regarding selection and formulation of the data inputs, choice of the algorithm and management of the available historical data.

Second, ML algorithms must be carefully controlled and validated to have out-of-sample efficacy. As one example, fitting a deliberately restricted version of an algorithm to many random selections of historical data may produce a more robust result than fitting an elaborate version to the full historical data sample. As another example, ML research often requires training several different versions of an algorithm and choosing among them. In this model selection step, researchers must be especially careful not to look ahead and consume data that should be set aside for true “out-of-sample” testing.

Third, while understanding the drivers of ML algorithm behavior can be a challenge, there may be means to gain insight. Application of ML doesn't automatically turn the investment process into an impenetrable black box. Decision trees, for example, allow for inspection of which predictors an algorithm deems most predictively valuable. We may also be able to assess a trained algorithm's likely response to circumstances of special interest by feeding it contrived data. As well, many longstanding research best practices remain applicable to ML contexts. These include comparing an ML-based forecaster's behavior to more transparent signals, analyzing whether it loads on known risk factors, and examining whether it shows signs of instability will induce portfolio turnover and increase costs.

Quantitative investing has always been about constant innovation, driven by adaptation to changes in the investing environment and the relentless pursuit of opportunities afforded by new insights, tools and resources. ML's integration into quantitative investing is an exciting frontier of that decades-long ongoing evolution. But its successful application must be informed by the special challenges posed by the market context, and beneficial research ML requires process discipline and puts a premium on investment domain knowledge. For asset owners, ML's flexibility and opacity will exacerbate challenges in strategy evaluation but working knowledge of ML approaches and ML research best practices offer a valuable defense.

Seth Weingram is senior vice president and director, client advisory, at Acadian Asset Management LLC, Boston. This content represents the views of the author. It was submitted and edited under P&I guidelines but is not a product of P&I's editorial team.

Disclosure & Disclaimer

This article has been reproduced and republished and is used with permission from Pensions & Investments. This material expresses the views and opinions of the authors and is being provided for informational purposes only. It should not be construed as investment advice, or an offer to sell or a solicitation of an offer to buy any security or as an endorsement, recommendation, or sponsorship of any company, security, index, strategy or advisory service, or to establish any investment relationship with Acadian or other investments managers mentioned. This material has not been updated since it was published and may not reflect the current views of the author(s) or recent market activity. No representation or warranty, express or implied, is made as to the accuracy, reliability or completeness of such information. Historical economic and performance information is not indicative of future results. The views expressed are subject to change based on market and other conditions.

This article may not be reproduced or disseminated in whole or part without the prior written consent of Pension & Investments.

© Copyright Pensions & Investments 2019