Machine learning revolution is still some way off

DAVID STEVENSON

Main Street and, let’s be honest, Wall Street love a great, revolutionary disruption narrative and in asset management there is no better hot theme than machine learning (ML) and artificial intelligence.

The robots are coming for fund manager jobs! As a journalist I, too, love these grand portrayals but I now realise that the largest changes happen at incremental level, where they creep up on an industry before turning into revolutionary value creation/destruction cycles.

My guess is that we are still early in the incremental phase with ML — in the “let’s see how it works” stage. And how is it working? A good place to start is to talk to those quants obsessed by factor-based investing, where large data sets exist and there are many competing narrative methodologies all tested to destruction by academics. The promise is obvious.

ML should be able to overcome the behavioural biases that still exist in factor-based investing despite the wall of data. Even here, though, you will find plenty of warning signs.

Peruse the pages of industry publications such as the Journal of Financial Data Science and you will find the likes of Joseph Simonian, director of quantitative research at Natixis, who observed...
(to Risk.net) that “many people on the Street think that to do financial data science, you just take a machine-learning algorithm and a data science model, and wholesale apply it to finance. But our argument, and the basis of this journal, is that financial data . . . has its own peculiarities.”

The challenge? Capital markets are constantly adaptive and data sets may not be deep and long enough to reveal any enduring trends. Commentators such as Rob Arnott of Research Affiliates and Campbell Harvey, a professor at the Fuqua School of Business in North Carolina, have already proposed a checklist for applying ML techniques with a focus on the depth of those data sources. In the absence of this, Mr Arnott suggested to Risk.net that using sparse data to train ML algorithms was akin to driving a Ferrari on a dirt track.

Andrew Lapthorne, head of quantitative equity research at Société Générale, probably feels a bit like that Ferrari’s driver. Using ideas articulated in a strategy paper in the summer of 2017, Lapthorne and his colleague Georgios Oikonomou have been running a live version of an ML strategy based on quant factors since the start of this year. It is early days but it is fair to characterise his early results as cautiously optimistic. What is much more interesting are his observations about how ML works in practice.

Let’s get some basics out of the way. The SocGen team looked at 80 factors, using a universe that spanned the period from December 2005 to January 2019, covering four regions: Asia ex Japan (MSCI World Asia ex Japan), Europe (Stoxx 600), Japan (Topix) and of course the US (the S&P 500 and S&P 400).

As to the mechanics of the ML system, the SocGen team used something called LightGBM, an open-source, implementation of an ML algorithm called a gradient boosting machine. The model tried to predict month-ahead performance but almost immediately the researchers discovered that the portfolios turned over crazily fast — a 50 per cent turnover rate meant the portfolio could look entirely different after just a few months. That high turnover suggests that trading costs could be a huge impediment to implementing the strategy.

Many enthusiasts for ML also love to conjure up scenarios where alternative data sources are incorporated. Mr Lapthorne, though, is less than convinced, warning that in financial markets “even a minor look-ahead bias in your data set will be picked up by the model leading to extraordinary returns”. So, the combination of new data and powerful data mining tools is a hazardous mix. His message? Stick with traditional fundamentals.

What this more conservative approach does reveal is that ML routines seem to learn from their own data and start reinterpreting time series in a dynamic way, with no look-ahead bias.

“The machine has no idea in 2007 that a financial crisis is on the way when it is building its model in 2006,” says Mr Lapthorne. “Equally, the machine can feed back its interpretation of our data to
help us understand our existing quant models. It is also excellent at recognising non-linear factors such as balance sheet risk.”

And here is one last crucial insight. It seems that ML routines can be especially useful on working out what to buy (and sell) in regime changes, ie, big shifts in interest rates. Unanchored to the past and with no look-ahead bias they become especially powerful, although that needs to be balanced by the observation that during periods of pronounced volatility the models don’t seem to keep up as well.

The upshot from Mr Lapthorne and Mr Oikonomou is, as I said, cautiously optimistic, with the models delivering about 7 per cent overall return with an investment return of 1.7 per cent since 2006.

“[After] trading costs, the alpha generated from factors in recent years is around 400 basis points. The big gains comes from a significant reduction in risk to under 4 per cent,” said Mr Lapthorne.

But even he seems far from convinced that there is big capacity in these strategies, which rather suggests to this cynical observer that ML is still very far from take-off.

If the factor quants ever so slightly struggle to make this work at scale, what hope is there for the rest of the industry?

Copyright The Financial Times Limited 2019. All rights reserved.