Machine Learning In Finance – 25 Years On

15 June 2015

In the early 1990s excitement grew around the potential application of machine learning and artificial intelligence techniques to finance. Take this introduction to a book in 1993:

"Two crucial developments occurred around 1980; both were enabled by the general availability of powerful computers that permitted much longer time series to be recorded, more complex algorithms to be applied to them, and the data and the results of these algorithms to be interactively visualized. The first development, state-space reconstruction by time-delay embedding, drew on ideas from differential topology and dynamical systems to provide a technique for recognizing when a time series has been generated by deterministic governing equations and, if so, for understanding the geometrical structure underlying the observed behaviour. The second development was the emergence of the field of machine learning, typified by neural networks, that can adaptively explore a large space of potential models. With the shift in artificial intelligence from rule-based methods towards data-driven methods, the field was ready to apply itself to time series, and time series, now recorded with orders of magnitude more data points than were available previously, were ready to be analysed with machine-learning techniques requiring relatively large data sets."

The quote is from the introduction to 'The Future of Time Series' (1993) by Neil A. Gershenfeld and Andreas S. Weigend, a book that reported on a competition set up by the Santa Fe Institute to explore the variety of new machine learning and other non-traditional forecasting methods that developed in the late 80s and 90s. Over the next decade many researchers pursued the research programme alluded to above.

Fast forward to 2015 and the investment community appears to be once again abuzz with talk of machine learning and artificial intelligence, as the next revolution in investment management! So what on earth happened in between, didn't we all do this before, 25 years ago?

Well, what happened was that despite the early promise, machine learning and AI did not turn out to be a game changer in the world of finance and economics. And this wasn't just confined to finance, many other areas such as image and speech recognition fell back on simpler approaches, and the highly rated tool of the time, the neural-network, slowly drifted out of the public conscious. The simplest explanation for this is that neural networks did not produce tangible improvements in understanding or prediction. Within finance applications, most economists were not in the least surprised, and could probably be paraphrased as saying "Well what did you expect? We told you markets are more or less efficient and impossible to predict. Building more complex tools is completely pointless". And even in tasks without such binding theoretical bounds on predictability, such as image recognition, it was soon discovered that although neural networks (and other associated techniques) could in theory solve the problems at hand, in practice it was extremely difficult to train them to do so. Instead most modelling and forecasting within finance continued via a process of researcher led model building, following largely manual processes of model selection and estimation. The recognition that a learning analogy might help didn't completely disappear though, and one could argue that Bayesian techniques that grew in popularity over the next two decades helped fill a gap.

So what's all the fuss about today in 2015? At the outset it must be acknowledged that the investment industry, particularly the quantitative end of it, has become increasingly commoditized and competitive in the last decade. And as fees have come down suppliers of investment management have looked to find ways to demonstrate they are different to their peers. Managers are splitting into two camps, those that say we will carry on doing what we've always done, and you can be rest assured our ongoing research will keep us competitive. And those that are claiming to do something genuinely different from the crowd. A cynic might argue that ML and AI are being used as a marketing trick by some in this latter camp.

But putting this cynicism aside, two things have patently happened. First, computing power has increased by a factor of 1,000s. Second, not everyone gave up on machine learning and some were quietly working away on improving algorithms and practical techniques that are today enabling machine learning to solve new sets of problems. Deep learning for example, one of the buzz terms of the moment, fits this story perfectly. With deep learning, thanks to some major breakthroughs in techniques for getting these networks to do what we want, people are finally starting to use the "deep" (multi-layer) neural networks that we knew were attractive in the 80s and 90s. Partly this breakthrough is processing power related as Moore's law has played out. But a big part of it is that humans have learnt that they need to take a much more hands on role in the process of training these systems. So this has led us to an almost paradoxical situation whereby it is simultaneously true that machines are able to do more and more tasks, better and better than they were able to before, but also humans are more deeply involved in that training process than they ever were before, and probably themselves ever expected or hoped to be. Machine learning works, but it really needs humans.

These two explanations for the renewed optimism still sound a lot like what people had said in the 1990s (see the opening quote), but there is however another development that supports the idea that this time it really is different – and that is "big data".

Through mass digitization and growth of the internet there has been a vast increase in the availability and dissemination of data. But big data brings with it big challenges and this is fuelling demand for techniques to filter out the good stuff from the irrelevant within these massive data sets. And this is half of what machine learning has always been about. One part of machine learning is about allowing algorithms to model data by evolving themselves over time. But the other defining aspect of machine learning is the ability for machines to discover relationships and associations within data sets that might have been previously obscured or not even conceived of by human researchers (generally what is known as "unsupervised learning"). This aspect of machine learning needs a lot of data by definition, because the whole idea is that the machine only has the data itself to go on. So the growth in big data and applications that stem from it (such as medical diagnoses, genetic research, internet shopping patterns) is stimulating improvements in machine learning.

So where does this lead us to in terms of machine learning and finance? Many of the reasons for the earlier disappointment have not gone away, and today's proponents of machine learning within investment management would do well to study the history of the last few decades experience before reinventing the wheel, or falling into known traps. The main point is that we have good reasons to believe that markets are, most of the time, pretty close to efficient, at least in a statistical sense that it should be extremely difficult to make consistent (risk-free) profits by forecasting the direction of prices. This is because by exploiting the information in the forecast (i.e. trading) the researcher/trader affects the future price. This reflexivity of markets makes the problem of learning about investing fundamentally different from one of learning to identify say a bicycle from a motor-bike in an image recognition task.

A second problem for applications in finance is the extremely limited history that machines can base their learning on. While it is possible, and common these days, to sub-sample history to the granularity of a micro-second, and this generates what looks and feels like unlimited time-series of data, it should be obvious that this does not necessarily add to our ability to detect features or create trading opportunities that can be exploited on the time scales that are useful to us. Someone who has used machine learning to identify a pattern of behaviour by web-site users that allows them to predict where a user is most likely to click next, can potentially monetise that across millions of other users almost instantaneously, and repeatedly. The cross-sectional and repetitive profit opportunities in those types of applications are huge. But most investing profit opportunities require the passing of time. Trading costs money and prices need to move sufficient distances before profits can be made.

Investing is also limited in that the cross-sectional opportunity set is relatively small. For example, anyone wanting to apply machine learning to futures markets trading should recognise that over any time interval there is a finite set of portfolios (limited by the product set) that can be built, and in practise a manager can only hold one portfolio at a time (even multi-strategy managers in effect hold a single net portfolio of positions at any point in time). Furthermore it is now increasingly recognised that over any time interval there are only a few distinct risk factors underlying movements in asset-prices, which taken together means the effective opportunity set is even more restricted than what the combinatorial possibilities a product set may appear to offer.

So the intrinsic reflexivity of markets, short data histories and limited opportunity sets pose major obstacles limiting the ability for machine learning to tangibly improve on other techniques. But there is an area in which machine learning might get to play a wider role in increasing our understanding of markets and investor behaviour, and that is the interface between machine learning and the study of how individuals themselves learn to collect, process and act on information.

Quietly but remarkably there has been something of a paradigm shift occurring within the economics profession over the same last 25 years, where nowadays it is perfectly acceptable to study how individuals form their expectations and make decisions. It may seem weird that this was never always the case, but for many years, for matters of both (convincing) theoretical argument and mathematical convenience, the fact that agents might behave in any way out of kilter with a rational utility maximising choice theory was largely ignored. But today learning itself has become part of the mainstream economics agenda, and this is cutting across both theoretical, empirical and experimental research. The study of computer learning algorithms is obviously very closely related.

Related to the study of learning is a body of research based on the building of artificial agent models (within computers) which offers an alternative method by which to study economic phenomena. For example scientists have simulated the interaction of many traders within artificial stock markets in order to get insights into how price dynamics such as bubbles and crashes might be generated. The "traders" in these systems are actually just algorithms, and in the more advanced approaches these algorithms learn how to trade over time. In other words some researchers are using artificial agent models, which in turn rely on machine learning, as a means by which to get insights into the methods and effects of human learning.

One obvious application of these techniques that could impact the investment management world is in the assessment of risk. A standard way of overcoming the short data histories we have in finance is to create artificial data on which to build and assess investment strategies. This principle is long established, and the calculation of value-at-risk, and stress-tests often relies on this basic idea. One can create future risk scenarios by randomly splicing together shorter samples of past data. One can ask therefore what might be the risks to my strategy if there is a 1998 style LTCM crisis immediately followed by a 2008 style surprise rate cut episode, even though this conjunction of events never occurred in history.

But creating scenarios from the same limited pool of historical data is clearly limited, so another approach is to build artificial agent models that can generate artificial, yet realistic data which can hopefully make up for the lack of short histories. In some sense one can see this is a kind of flight-simulator approach. In reality a pilot cannot (and would not want to!) face a crash landing situation many times, but in a computer simulation they can expose themselves to risks and learn how to behave without actually risking anyone's life. A flight simulator can create risks that may have never actually occurred, for example flying into a volcanic dust cloud in the middle of a thunder-storm. A generative model of risk could create this scenario even though it may have never occurred in history.

In fact it is precisely this ability to use machine learning based systems to generate data, rather than just analyse it, that has driven some of the key technological advances that are underpinning the advances in deep learning that is fuelling the hype in finance today. One of the world's leading machine learning experts Geoffrey Hinton developed this idea and summarised it succinctly in the title of a now seminal paper, "To recognize shapes, first learn to generate images" (Hinton, 2007).

Lessons from machine learning can make these risk simulations even more lifelike, particularly in the economic field where agents actions directly influence and alter the environment they are participating in. In these situations it is best to think of individuals learning (and acting) within an environment which is composed of many other individuals also learning about the environment. Think of Keynes' famous beauty contest, but in a dynamic context. The critical input from machine learning is that a trading strategy can be tested in a strategic reflexive scenario, where the risk scenario itself will endogenously evolve in response to the trading strategy. Researchers and regulators are using these artificial agent techniques to study how high-frequency algorithmic trading techniques could be a source of instability in markets e.g. Hayes et al (2012).

But while there may or may not be improved performance for investors via the use of machine learning, there will certainly be increasing challenges for regulators, managers and investors alike.

Take for example the thorny issue of what constitutes a "trading error". Investors these days often try to ensure that managers compensate them for any losses caused by errors. But what is an error when the trading system is following an AI inspired machine learning approach? Unprofitable trades will be made. The system may behave in a way that was unanticipated at the outset by both the manager and the investor. The machine may generate not just trades at a faster speed than a human can oversee (as algorithmic systems currently already do), but also generate new algorithms at a faster rate than a human can oversee. All these abilities could be argued to define precisely the key benefits of a machine learning approach, so there is perhaps something intrinsically difficult about the challenge of reconciling machine learning and human oversight.

Machine driven systems will also need to be aware of regulatory restrictions and work within them. In other fields of learning one would think it acceptable to learn by "dipping one's toe in the water", literally as one tries to learn the temperature of water. But what would be an equivalent action by a trading machine learning algo? An explorative algo might want to submit an order to a market in order to test the liquidity or resilience of the market. The market response may cause it to cancel the order. There may have been no intention by the algo to commit to trade the full amount posted, in the same way the swimmer is not committing to diving in the sea when they dip their toe in the water. Once the market has responded the algo may spot an opportunity to trade on the other side of the market. Is this spoofing?

And how will machines reflect a more general sense of ethical conduct that underpins regulatory frameworks worldwide? Machines are in theory fantastic at finding solutions by developing rules of behaviour within a set of preordained rules or a well-defined framework – that is almost by construction exactly what ML is designed to do. But what if the solutions found are deemed by humans to be within the rules but against the spirit of the rules? These situations arise all the time in the real world, and humans can generally understand the difference between the two, though sometimes a court is needed to help them along. There are real challenges here, and the one thing that it is safe to predict is that these issues will surely be the cause of some attention among risk managers, compliance and regulators, Paddrik (2013). The perceived computational complexity of machine learning will also not help the situation, and so my sense is that some protocols and broad principles will need to be defined rather quickly that address this coming issue.

In summary, although in some ways we have been here before there are some genuinely new factors today that does mean some of the buzz is warranted. Partly this is related to the challenges and opportunities created by big data. But I suspect an ultimately more profound impact will come from a more general advance in our understanding of individual behaviour, through the use of more computational approaches to research, within which machine learning will play a key role. We will be forced to ask legal, ethical and even philosophical questions along the way too.

References

Neil A. Gershenfeld and Andreas S. Weigend, 'The Future of Time Series' (1993)
http://www.santafe.edu/media/workingpapers/93-08-053.pdf

Hinton, G. E., 'To recognize shapes, first learn to generate images', (2007) in P. Cisek, T. Drew and J. Kalaska (Eds.) Computational Neuroscience: Theoretical Insights into Brain Function. Elsevier.
http://www.cs.toronto.edu/~hinton/

Hayes, Roy and Paddrik, Mark E. and Todd, Andrew and Yang, Steve Y. and Beling, Peter and Scherer, William, Agent Based Model of the E-Mini S&P 500 Future: Application for Policy Making (2012). Proceedings of the 2012 Winter Simulation Conference. Available at SSRN:
http://ssrn.com/abstract=2112139 or http://dx.doi.org/10.2139/ssrn.2112139

Paddrik, Mark E., Assessing Financial Markets Through System Complexity Management (December 2013). Available at SSRN:
http://ssrn.com/abstract=2400998 or http://dx.doi.org/10.2139/ssrn.2400998

Disclaimer

This research note is published by Neuron Advisers LLP ('Neuron'). You have been provided with this material only upon your acceptance of these Terms and Conditions herein.

Neuron Advisers LLP is authorised and regulated by the Financial Conduct Authority, with firm reference number 563919. Neuron Advisers LLP is registered with the Commodities Futures Trading Commission under NFA ID 0439462.

Investments or investment services mentioned in this research note are not available for investment and are not being marketed in any jurisdiction. The contents of this research note are not intended to be read by any persons in any jurisdiction other than the United Kingdom.

This material is presented for information purposes only. It is intended for your personal, non-commercial use. No information or opinions contained in this material constitute a solicitation or offer by Neuron to buy or sell securities or to furnish any investment advice or service. Neuron does not provide investment advice, tax advice or legal advice through this material and you agree that this material will not be used by you for such purposes.

This material is intended as a general introduction to Neuron and the Neuron research blog by which means Neuron can express ideas and opinions. The material contained herein is the sole opinion of Neuron.

The information provided in this research note is intended for institutional investors and Professional Clients and Eligible Counterparties as defined by The Financial Conduct Authority and for those who are considered as qualified eligible persons as defined by Commodities Futures Trading Commission Regulation 4.7. It is not intended for retail investors.

The contents of this research are not intended for distribution to, or use by, any individual or entity in any jurisdiction where their distribution or use would be contrary to local law or regulation or which would subject Neuron Advisers LLP to registration with the jurisdiction. You should be aware that any rules and/or regulations applicable to providing financial services (and the resultant investor protections that may be available), may not apply to persons who obtain information from the internet and its various applications, of which this material forms part.

Neuron Advisers LLP assumes no responsibility for access to this material by any person located within a country or jurisdiction where such access would be contrary to any law or regulation in that country.

We try to ensure that the information in this research note is correct, but we do not give any express or implied warranty as to its accuracy, timeliness or completeness, nor is Neuron under any obligation to update such information. Any data supplied has not been audited and is provided for information purposes only.

We are not liable for any damages (including, without limitation, damages for loss of business or loss of profits) arising in contract, tort or otherwise from the use of or inability to use this research note, or any material contained in it, or from any action or decision taken as a result of using this research or any such material.

This research note may provide links to other research and websites. We do not control the linked sites or research and we are not responsible for the contents of any linked site or research, any link in a linked site or research, or any changes or updates to such sites or research. We provide these links to you only as a convenience, and the inclusion of any link does not imply our endorsement of the site or research.

Neuron Advisers LLP is a limited liability partnership registered in England & Wales with number OC367248 and registered office address at 2nd Floor 11 Albemarle Street London W1S 4HH.

Any questions about the contents of this material should be directed to: enquiries@neuronadvisers.com.