Financial Data Scientist – a new investment profession

Ralf Frank, DVFA, Frankfurt/Main
Prof. Dr. Andreas Hoepner, UCD, Dublin
Dr. Damian Borth, German Research Center for Artificial Intelligence (DFKI), Kaiserlautern

A Google search returns roughly 373,000,000 hits on the term ‘Big Data’, and around 73,500,000 for ‘Big Data + Finance’. A search for ‘Financial Data Science’ actually returns as many as 56,800,000 entries. But, what is meant exactly by Financial Data Science? Is it just another buzzword in the context of the much-touted *digitalization* of the economy?

When artificial intelligence is mentioned in the financial market, then many immediately imagine losing jobs they fear will be replaced by algorithms or robots. In truth, Artificial Intelligence (AI), will end up creating thousands of jobs in areas such as design, ‘feeding’ and programming. Such new jobs are already being created, numbering in the tens or even hundreds of thousands – not only in the financial sector, but in society as a whole. But this new job creation comes with changing demands on skills and qualifications.

Artificial Intelligence is the stuff of Hollywood storytelling. The real change will come from Augmented Intelligence.

As an acronym, AI can have two different meanings: the well-known ‘Artificial Intelligence’ and the as yet little-used ‘Augmented Intelligence’. While the first is a concept often found in science fiction ("Who is more intelligent: man or machine?"), Augmented Intelligence is about intelligent machines in interaction with human experts, who benefit from better performance in less time through the coupling of man-machine. By itself, Artificial Intelligence is used in areas where a routine task, such as the recognition of patterns, data or people, can be repetitively performed by a machine more precisely and with fewer errors than by humans. One example is passport control at airports using biometric data.

But, despite the palpable hype surrounding Artificial Intelligence: the future lies in Augmented Intelligence – especially for the financial market. In cases where pattern or image recognition doesn't have to run in real time, and where flexibility, rather than dull repetition is required, AI's programming and learning processes are either too expensive or too inflexible. People can thus perform certain tasks using an Augmented Intelligence process with comparable accuracy but greater flexibility.

Self-learning algorithms are a metaphor. Predictions, e.g. by neural networks, have weaknesses that require human intervention.

Algorithms, it is asserted time and again, can learn. But when it comes to tasks that do not involve things like recognition of a face in a passport, but rather the prediction of an unobservable result, three major weaknesses of the Deep Learning concept emerge:

1. The results that a neural network programmed on the basis of Deep Learning can provide depend largely on the weighting of factors attached to the input data. In contrast to conventional statistics, they cannot integrate a measure of uncertainty, such as confidence intervals.
2. Deep learning could actually be called Deep *Imitation*, because if algorithms learn, they are merely imitating. We are still a long way off from the development of an independent understanding of interrelationships or an independent knowledge base – even 20 years after HAL 9000, the omniscient computer from Stanley Kubrick's "2001 - A Space Odyssey", algorithms are not there yet.

3. Most deep learning neural networks are so complex that even the scientists who program them usually don't understand exactly how decisions or calculations are made in the deeper layers of the network.

Taken together, these three weaknesses illustrate why, even in 2018, we are still far from the ability to leave algorithms or computers on their own to create forecasts whose probability or accuracy cannot easily be verified. This also explains why, in view of imperfect technology, we will continue to create tens of thousands of jobs in the future, especially in the financial market, the core of which is forecasting: in order to be able to ensure the effective and targeted use of this technology, whose advance can no longer be stopped.

As long as the customer is a human being, thousands of jobs will be created in the translation of algorithmic outputs into human logic.

In investment management, active investment is increasingly on the defensive. "Quants" have long had a place in the financial markets, but it seems that the golden age for quantitatively oriented investment professionals has only just begun. Despite this optimism, we advocate a proactive approach to the issues of digitalisation: First, the current upheavals in the financial industry offer an ideal point to rethink methods and investment approaches. Second, investment professionals should be working to develop their own digitalisation skills.

Algorithms and neural networks need to be designed and programmed for financial markets. It is wrong to think that this task can only be accomplished by programmers and IT specialists. Two trends could converge here: MiFID II will presumably lead to reduced investment research capacities over the coming months and years, which can no longer be commercially exploited under the framework of conventional research. It is unacceptable that this know-how should leave the marketplace.

Among the investment professionals who will be affected by this are many experts whose loss would be deplorable. In view of our thesis that the future will be shaped by an intelligent interaction between humans and machines, many of today's financial analysts and fund managers could represent a new generation of future investment professionals, namely Financial Data Scientists who work in information identification (What information? How to weight it?) and make their contribution, e.g. in testing AI. However, this will require new skills and abilities.

A small team of five Financial Data Scientists with statistical expertise and programming skills can be very competitive compared to five portfolio managers and two dozen analysts working on outdated technology that hasn't been updated since the last century.

In their own interest, investment professionals should consider digital technologies, especially AI, to be tools and methods. Freely available, yet anything but trivial, programming languages like R or Python are the minimum standard in Financial Data Science today. To quote from a recent employment ad for a large global asset manager: "We offer new products that provide access to our data inventory for data scientists, machine learning researchers and quantitative model builders... our target customers use the R statistical computing language and apply mathematical and statistical methods to investment and risk issues... our platform allows them to apply their algorithms to our databases, and not vice versa. Requirements are... experience with R, experience with Python, a great interest in the data that underpins our financial system."

Whoever deals with R or Python will recognise the power these programs provide for the analysis of a wide variety of data sets. Analysis with Python is to conventional analysis instruments what the Internet is to card index boxes. Already, Master of Finance graduates

entering the financial job market today from leading universities, usually arrive with experience in programming Python or R. Bloomberg now offers an interface to Python - as a logical alternative to the conventional Excel interface. Knowledge of Python and R should be standards for investment professionals, which is why training in this area, such as the Chartered Financial Data Scientist programme at DVFA (http://bit.ly/2p09LAJ), rely on Python.

This is especially important since FinTech companies whose services are often based on these technologies are slowly but surely establishing themselves as market players. Smart Beta became Factor Investing, which is now developing to become *evidence-based investing* with the help of Financial Data Science. Which client, in 2025, will still be interested in investment strategies without transparent, scientifically sound evidence?

By now, everyone in the market should be aware that practically every major asset manager today already maintains teams of Financial Data Scientists, or are building such teams as we speak. What are investment professionals waiting for?

The German version of this article was published in Börsen-Zeitung on 08 March 2018.