Artificial Intelligence and The Future of Finance

Miquel Noguer i Alonso PhD

IEF, Columbia University and ESADE

Abstract

We provide in this paper an analysis of how Artificial Intelligence (AI) companies are and will be changing the Financial Services Industry. In particular two main areas will be disrupted by AI and Machine Learning (ML) technologies: Financial Marketing and Investing.

Both disciplines are in the middle of a big data revolution: larger, structured and unstructured datasets, modern Data Analysis that makes extensive use of Ai and ML and faster and cheaper computing paradigms.

Financial Marketing is now digital marketing and that means the inevitability of algorithms taking over the value chain process. Consumers want a 24 / 7 easy access to services and products that can be achieved only obviously with a high degree of automation. These are uncharted territories algorithms should help to customize the new digital experience.

Investing is an information processing game, AI and ML will help there in modeling unstructured data using Natural language Processing techniques to measure for example sentiment and some Machine Learning supervised and unsupervised techniques that allow for more flexibility in capturing highly dimensional non-linear data generating processes. All models in financial markets obviously should deal with very noisy data generating processes possibly non-stationary.

All financial services firms should expand their technology / modeling toolbox to these techniques. They will transform finance the only open question is when.

Revolutions in Finance

Financial Services Industry is living an unprecedented structural change. There are several technologies that are transforming all different sectors of financial services industry:

- **Digitalization.** The sector is still figuring-out how to service clients in the digital ecosystem. How to deal with digital clients, retain and gain market share. Financial firms are making significant investments to enhance customer experience and engagement through the development of new digital products and capabilities. Throughout the industry, there is widespread digital disruption. Digital technologies are disrupting the industry to a great or moderate extent.
  This is uncharted territories for all industries as marketing, big data department are trying to understand the new challenges and opportunities.
- **Big Data and Artificial Intelligence.** We will focus in this paper in the growing importance of artificial intelligence ad machine learning applications in the industry. Modern data analysis will make an intensive use of artificial intelligence and machine learning models to use large internal and external datasets
- **Blockchain**. The blockchain technology, which is a digital, distributed transaction ledger with identical copies maintained on each of the network's members' computers. All parties can review previous entries and record new ones. Transactions are grouped in blocks, recorded one after the other in a chain of blocks (the 'blockchain'). The links between blocks and their content are protected by cryptography, so previous transactions cannot be destroyed or forged. This means that the ledger and the transaction network are trusted without a central authority – and intermediaries.
- **Internet of Things**. Many industries see great promise in IoT applications: Analysts and technology providers forecast added economic value of anywhere from $300 billion to $15 trillion by this decade's end. Few analysts seem to expect anything but a transformative effect on just about every dimension of economic activity by 2020. Some use cases have already proven themselves: Applications such as auto insurance telematics and "smart" commercial real estate building-management systems offer clear IoT examples of new products or changed processes.
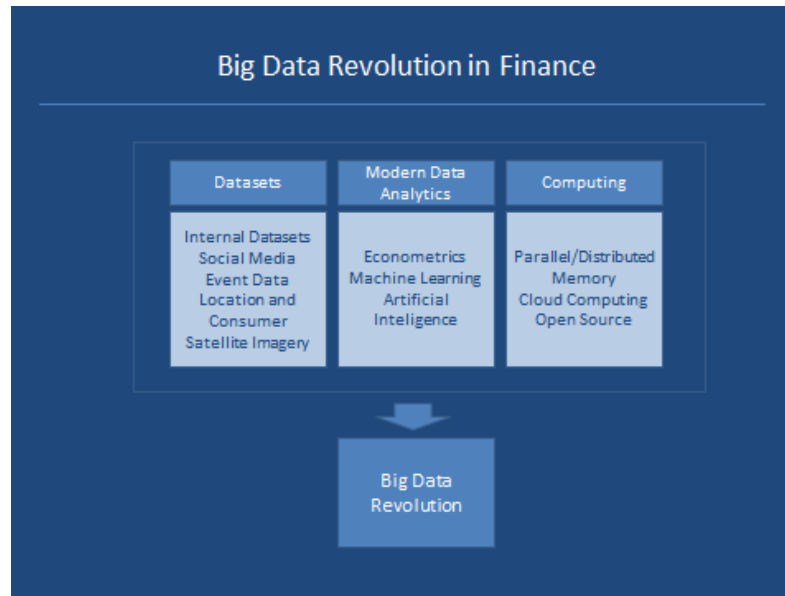
Big Data Revolution in Finance

Figure 1: Big Data Revolution in Finance

There have been three main drivers the Big Data revolution in finance

1) Datasets: Exponential increase in amount of datasets (internal and external information)

2) Computing: Increase in computing power and data storage capacity, at reduced cost

3) Modern Data Analytics: Artificial Intelligence and Machine Learning methods to analyze complex datasets

Datasets

Exponential increase in amount of data: With the amount of published information and collected data rising exponentially, it is now estimated that 90% of the data in the world today has been created in the past two years alone. This data flood is expected to increase the accumulated digital universe of data from 4.4 zettabytes (or trillion gigabytes) in late 2015 to 44 zettabytes by 2020. The Internet of Things (IoT) phenomenon driven by the embedding of networked sensors into home appliances, collection of data through sensors embedded in smart phones (with ~1.4 billion units shipped in 2016), and reduction of cost in satellite technologies lends support to further acceleration in the collection of large, new alternative data sources.

Financial service companies need to figure out how to use internal and external databases to do marketing and also for investing purposes in asset managers.

Computing

Increases in computing power and storage capacity: The benefit of parallel/distributed computing and increased storage capacity has been made available through remote, shared access to these resources. This development is also referred to as Cloud Computing. It is now estimated that by 2020, over one-third of all data will either live in or pass through the cloud. A single web search on Google is said to be answered through coordination across ~1000 computers. Open source frameworks for distributed cluster computing (i.e. splitting a complex task across multiple machines and aggregating the results) such as Apache Spark have become more popular, even as technology vendors provide remote access classified into Software-as-a-service (SaaS), Platform-as-a-service (PaaS) or Infrastructure-as-a-service (IaaS) categories. Such shared access of remotely placed resources has dramatically diminished the barriers to entry for accomplishing large-scale data processing and analytics, thus opening up big/alternative data based strategies to a wide group of both fundamental and quantitative investors.

Modern Data Analytics

Machine Learning methods to analyze large and complex datasets: There have been significant developments in the field of pattern recognition and function approximation (uncovering relationship between variables). These analytical methods are known as 'Machine Learning' and are part of the broader disciplines of Statistics and Computer Science. Machine Learning techniques enable analysis of large and unstructured datasets and construction of trading strategies. In addition to methods of Classical Machine Learning (that can be thought of as advanced Statistics), there is an increased

Artificial Intelligence Definition

Artificial intelligence (AI, also machine intelligence, MI) is intelligence demonstrated by machines, in contrast to the natural intelligence (NI) displayed by humans and other animals. In computer science AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals. Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving".

The scope of AI is disputed: as machines become increasingly capable, tasks considered as requiring "intelligence" are often removed from the definition, a phenomenon known as the AI effect, leading to the quip "AI is whatever hasn't been done yet." For instance, optical character recognition is frequently excluded from "artificial intelligence", having become a routine technology. Capabilities generally classified as AI as of 2017 include successfully understanding human speech, competing at the highest level in strategic game systems (such as chess and other games), autonomous cars, intelligent routing in content delivery network and military simulations.
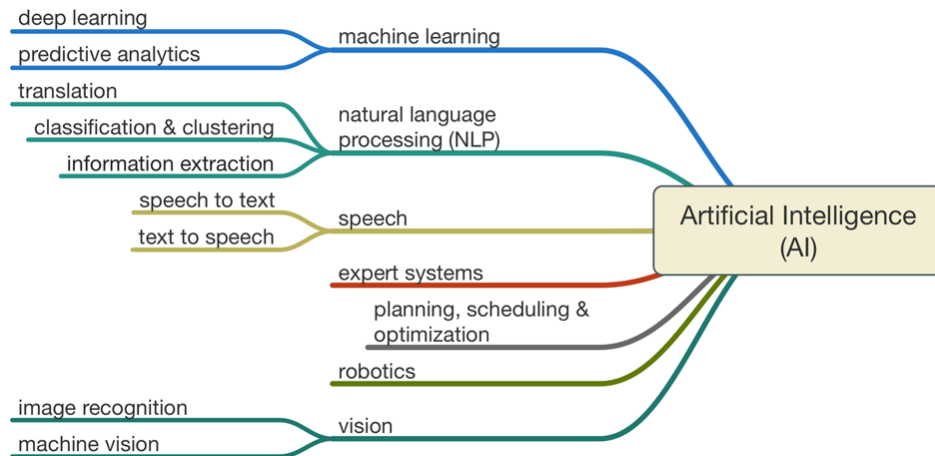
Artificial intelligence was founded as an academic discipline in 1956, and in the years since has experienced several waves of optimism, followed by disappointment and the loss of funding (known as an "AI winter"), followed by new approaches, success and renewed funding. For most of its history, AI research has been divided into subfields that often fail to communicate with each other. These sub-fields are based on technical considerations, such as particular goals (e.g. "robotics" or "machine learning"), the use of particular tools ("logic" or "neural networks"), or deep philosophical differences. Subfields have also been based on social factors (particular institutions or the work of particular researchers).

The traditional problems (or goals) of AI research include reasoning, knowledge, planning, learning, natural language processing, perception and the ability to move and manipulate objects. General intelligence is among the field's long-term goals.

Approaches include statistical methods, computational intelligence, and traditional symbolic AI. Many tools are used in AI, including versions of search and mathematical optimization, neural networks and methods based on statistics, probability and economics. The AI field draws upon computer science, mathematics, psychology, linguistics, philosophy and many others.

First coined in 1956 by John McCarthy, AI involves machines that can perform tasks that are characteristic of human intelligence. While this is rather general, it includes things like planning, understanding language, recognizing objects and sounds, learning, and problem solving.

Arthur Samuel coined the phrase not too long after AI, in 1959, defining it as, **"the ability to learn without being explicitly programmed."** You see, you can get AI without using machine learning, but this would require building millions of lines of codes with complex rules and decision-trees.

We can put AI in two categories, general and narrow. General AI would have all of the characteristics of human intelligence, including the capacities mentioned above. Narrow AI exhibits some facet(s) of

human intelligence, and can do that facet extremely well, but is lacking in other areas. A machine that's great at recognizing images, but nothing else, would be an example of narrow AI.
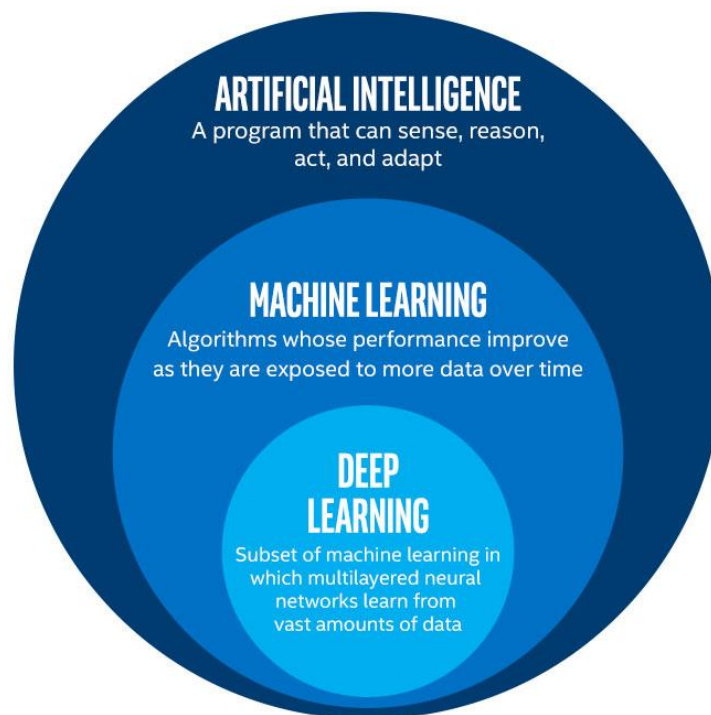
Figure 3: Artificial Intelligence versus Machine Learning versus Deep Learning

Machine Learning definition

Machine learning is a field of computer science that gives computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.

The name Machine learning was coined in 1959 by Arthur Samuel. Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data – such algorithms overcome following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs. Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms with good performance is difficult or infeasible; example applications include email filtering, detection of network intruders or malicious insiders working towards a data breach, optical character recognition (OCR), learning to rank, and computer vision.

Machine learning is closely related to (and often overlaps with) computational statistics, which also focuses on prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field. Machine learning is

sometimes conflated with data mining, where the latter subfield focuses more on exploratory data analysis and is known as unsupervised learning. Machine learning can also be unsupervised and be used to learn and establish baseline behavioral profiles for various entities and then used to find meaningful anomalies.

Within the field of data analytics, machine learning is a method used to devise complex models and algorithms that lend themselves to prediction; in commercial use, this is known as predictive analytics. These analytical models allow researchers, data scientists, engineers, and analysts to "produce reliable, repeatable decisions and results" and uncover "hidden insights" through learning from historical relationships and trends in the data.

Effective machine learning is difficult because finding patterns is hard and often not enough training data are available; as a result, machine-learning programs often fail to deliver.

Let's define some important terms:

- Labeled data: Data consisting of a set of training examples, where each example is a pair consisting of an input and a desired output value (also called the supervisory signal, labels, etc)
- Classification: The goal is to predict discrete values, e.g. {1,0}, {True, False}, {spam, not spam}.
- Regression: The goal is to predict continuous values, e.g. home prices.

Types of machine learning Algorithms

There some variations of how to define the types of Machine Learning Algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following:

- Supervised learning
- Unsupervised Learning
- Semi-supervised Learning
- Reinforcement Learning

Supervised Learning

We can think of supervised learning as function approximation, where basically we train an algorithm and in the end of the process we pick the function that best describes the input data, the one that for a given X makes the best estimation of y (X -> y). Most of the time we are not able to figure out the true function that always make the correct predictions and other reason is that the algorithm rely upon an assumption made by humans about how the computer should learn and this assumptions introduce a bias, Bias is topic I'll explain in another post.

Here the human experts acts as the teacher where we feed the computer with training data containing the input/predictors and we show it the correct answers (output) and from the data the computer should be able to learn the patterns.

Supervised learning algorithms try to model relationships and dependencies between the target prediction output and the input features such that we can predict the output values for new data based on those relationships which it learned from the previous data sets.

The main types of supervised learning problems include regression and classification problems

List of Common Algorithms

- Nearest Neighbor
- Naive Bayes
- Decision Trees
- Linear Regression
- Support Vector Machines (SVM)
- Neural Networks

Unsupervised Learning

The computer is trained with unlabeled data.

Here there's no teacher at all, actually the computer might be able to teach you new things after it learns patterns in data, these algorithms a particularly useful in cases where the human expert doesn't know what to look for in the data.

This family of machine learning algorithms is mainly used in pattern detection and descriptive modeling. However, there are no output categories or labels here based on which the algorithm can try to model relationships. These algorithms try to use techniques on the input data to mine for rules, detect patterns, and summarize and group the data points which help in deriving meaningful insights and describe the data better to the users.

List of Common Algorithms

- Hidden Markov Models
- Principal Component Analysis
- K-Means

Natural language processing

To an extent, sophisticated analytics programs can help businesses utilize their data by searching for and revealing patterns hidden in structured data, such as spreadsheets and relational databases. But these sources only account for 20 percent of all available data. The real challenge for enterprises is getting value from a sea of social media posts, images, email, text messages, audio files, Word documents, PDFs and other sources that make up the other 80 percent of data that can't be understood by computers—information otherwise known as unstructured data. To extract value from unstructured data, companies across industries are turning to Natural Language Processing (NLP).

NLP enables computer programs to understand unstructured text by using machine learning and artificial intelligence to make inferences and provide context to language, just as human brains do. It is a tool for uncovering and analyzing the "signals" buried in unstructured data. Companies can then gain a deeper understanding of public perception around their products, services and brand—as well as those of their competitors.

A growing number of businesses are now using NLP. In fact, a recent report by research firm Markets and Markets predicts the NLP market will reach $13.4 billion by 2020, a compound annual growth rate of 18.4 percent.

Financial organizations in particular have applied NLP to obtain actionable insights from digital news sources. For example, NLP can help:

- Gather real-time intelligence on specific stocks. A financial services provider can write a query using Data News API that will create a real-time alert if analysts upgrade or downgrade a stock, thereby offering clients a valuable trading edge.
- Provide key hire alerts. Everyone knows when a large enterprise hires a new CEO, but company share price and performance can also be affected by the hiring (or departure) of talented marketing, sales, development and finance executives. The News API can deliver instant alerts regarding crucial changes to a company's management.
- Monitor company sentiment. While major news, such as an earnings report or an acquisition, affect how investors view a company, so too can the general tone of the news coverage. Using NLP tools financial services providers can track mentions of companies and discern negative or positive sentiment in news coverage.
- Anticipate client concerns. Banks and other financial institutions can use NLP to discover and parse customer sentiment by monitoring social media and analyzing conversations about their services and policies.
- Upgrade quality of analyst reporting. With the ability to access relevant, filtered information, financial services analysts are able to write more detailed reports and provide better advice to clients and internal decision makers.
- Understand and respond to news events. Individual companies and industries as a whole are subject to national and global events and government or judicial decisions. Using API's, financial

services companies can monitor news about, say, an oil spill for clients with holdings in that industry.

- Detect insider trading. Financial services providers can use NLP queries to track the incremental sale and purchase of insider shares that otherwise might fly under the radar. The financial services industry is dependent not only on information, but also on information delivered as close to real-time as possible. To access and analyze relevant information in the rapidly expanding universe of unstructured data, financial services providers are turning to natural language processing to help them make decisions and provide sound advice and quality products to clients.

Artificial Intelligence Applications in Finance

There are two obvious fields of application: Investing and Marketing.

Digital marketing and AI

1. Personalize user experience

It is the most crucial area where AI can break in and create a significant impact. A customer is the King of any business and content is king for a marketer. If he can align the content marketing strategy with artificial intelligence, it could be ground-breaking. Based on data collected such as customer searches, buying behavior and interests, customized content campaigns can be conducted. The catch is that it can be done for every single customer or prospect.

Chatbots are another example of AI interference in enhancing user experience. Chatbots are programmed to interact with customers on the basis of data that it receives. Traditional chatbox and text communication will soon give away to a multi-dimensional communication system with sensory abilities such as voice and touch. This would personalize the whole experience for users as they receive the impression of talking to a real person with consciousness.

Augmented reality, another aspect of AI can be leveraged to provide consumers an option to see and feel the product before actual purchase. This would make decision making easy for customers as they are able to perceive the product even before purchasing. This will stimulate a faster response from the customer and in turn increase the revenue.

2. Predictive Marketing

Each time a user browses on the internet; new data is generated and collected for AI analysis. This data can reveal information such as user needs, behaviors, and future actions. Based on this information, marketing can be optimized to supply the most relevant information. Social media outreach also reveals personal information about the prospect making it easy for marketers to create a targeted campaign.

This further reduces the sales-cycle as the relevant information is handed over to customers on a silver plate. This "predictive" campaigns can significantly reduce customer research on the product and makes

decision making easier. As for marketers, they can continue to analyze the buyer through data and even make the customer return!

This type of AI-enabled algorithms will challenge the current 'hotshots' like SEO in a huge way. With AI powering the digital marketing initiatives from ground level, chances of trends such as SEO, banner ads becoming obsolete are high. After all, who needs SEO and website traffic when you have a detailed report of your prospect?

3. Image recognition to increase ROI

Previously, image recognition was confined to identifying isolated objects in an image. But with AI – enabled software, it is now possible to get a detailed description of an image. Amazon's latest brainchild Amazon Rekognition can actually recognize human faces, emotions involved and identify objects.

This technology can be used in various ways for various sectors. For banking and financial sector, AI-enabled image recognition can be leveraged for faster payment processes and enhance customer security. Social media is a huge source of images.

Social media has always been biased towards visual content as tweets with images receive 150 more tweets and Facebook posts with images receive 2.3X times more engagement. The world population together shares 3.25 billion photos a day according to this research. This humongous amount of images can be leveraged by AI to understand consumer patterns, behaviors, and needs. AI software will look for images in social media and compare it to a large image library to draw conclusions.

For example, a snack manufacturer can map their brand against the huge collection of photos in social media and understand the buyer demographics such as age group, gender etc., and also geographical potentials such as if the snack was consumed more at the beach, park, supermarkets, theaters and so on. This will help to align the marketing strategies in order to extract maximum Return on Investment.

Artificial Intelligence Companies

We provide in figure 4 a list of artificial intelligence companies that can help financial services companies to explore, research and use these technologies in their businesses.
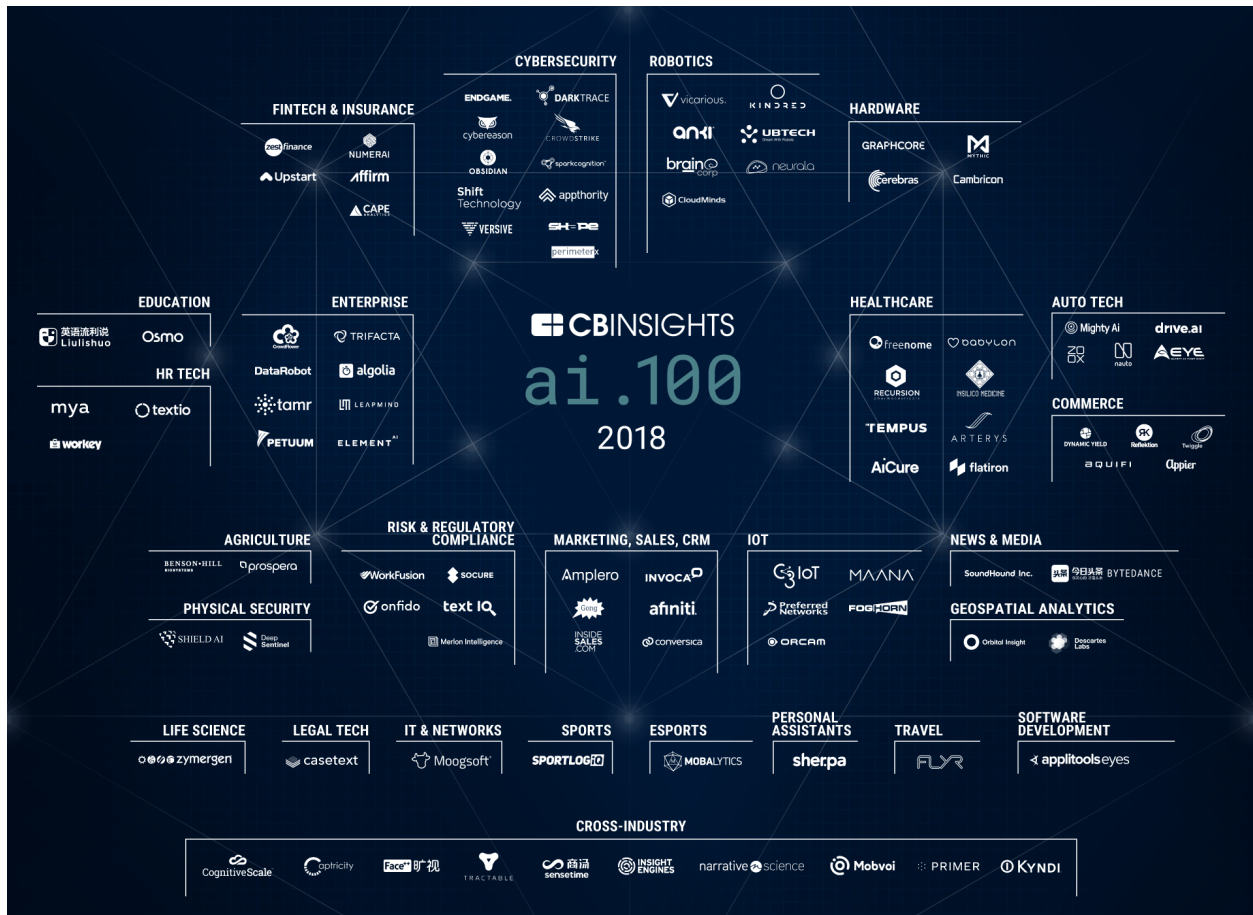
Investing and AI

All areas of investing can benefit from the use of AI and machine learning in addition to the use of completely new structured and unstructured datasets. We will provide in another paper more details of these advances. Here we provide a high level summary of the main applications.

Artificial Intelligence and machine learning can be useful in many investing applications:

- Using machine learning classification and regression techniques to select stocks, asset allocation. Boosting techniques like XG Boost and Adaboost are especially useful.
- Deep learning models to capture time series non-linearity
- Clustering techniques like PCA to reduce dimensionality and improve estimation.
- Natural Language Processing techniques sentiment analysis to incorporate in stock picking or asset allocation applications.

AI benefits

Main benefits are the following:

- Digest non structured data
- Capture non-linearities in the data
- Model flexibility

AI drawbacks

AI and ML algorithms operate in a very random, possibly non-stationary, highly dimensional when it comes to investing

- Non-stationarity
- Estimation / Learning
- Data snooping
- Transparency/Interpretation might be an issue

Conclusions

Artificial Intelligence is going to have a profound impact on finance, investing and marketing are one of the fields on which short and medium term high impact in the value chain.

Financial marketing, in the past being a human task is and will be replaced by smarter and smarter algorithms, as finance is going almost fully digital humans are being replaced by big data and algorithms. This is uncharted territory financial services firms that have on their data sets more information than ever have to figure out how to implement digital marketing properly. Customization is the way to go as al clients have different product preferences and level of service preferences. Algorithm should then play that role, much more research is needed to understand these new client digital ecosystem.

Investing is being disrupted by the big data revolution with new data sets, new models (mainly Artificial Intelligence / Machine Learning) and a new computing paradigm with the cloud and open source models. Investing is an information processing game, AI and ML will help there in modeling unstructured data using Natural language Processing techniques to measure for example sentiment and some Machine Learning supervised techniques that allow for more flexibility in capturing highly dimensional non-linear data generating processes, unsupervised to model "hidden" features. All models in financial markets investing obviously should deal with very noisy data generating processes possibly non-stationary.

The trend will continue in different directions, ultimately as AI gets better and better we will see a replacement of human tasks for AI's. The big question is not if the question is when.

Are you ready?

Dr. Miquel Noguer i Alonso

IEF, Columbia University i ESADE