Chapter 1

The Intelligence Revolution

What Is Today's Artificial Intelligence?

No matter how closely one works with technology, the discussion on a perceived imminent Artificial Intelligence revolution, as well as the prophetic brainstorming that takes place globally around what it brings, is impossible to escape us. A big part of this discussion is occupied by the debate around the social and economic implications of the new astonishing technological achievements; achievements that few people would deem possible just a few years ago and might cause a productivity explosion together with a labor market disruption a few years from now.

Artificial Intelligence started in the decade of 1950, as the quest for building machines capable of thinking like humans. Throughout the years, a number of schools of thought on AI formed, proposing different methodologies. None however have proved successful, never mind at achieving an ambitious goal such as human-level intelligence. Hence, the field has undergone prolonged periods of decline, known as "AI winters".

Very recently Machine Learning, has risen to become the absolute dominant approach to AI and its success is largely the reason why this conversation is taking place now.

Machine Learning has not achieved the original vision of intelligent machines, however, it already thrives in an extremely diverse array of important applications that we now use seamlessly in our daily lives. With the success of Machine Learning, the term Artificial Intelligence gradually shifted to mean something humbler than the original vision.

We now distinguish this original vision explicitly by calling it Artificial General Intelligence (AGI) or Strong AI. At the same time, in many cases and because of its success, Machine Learning is now sometimes used synonymously to AI bearing, its narrower definition: the array of applications that it powers extremely successfully, which is very wide but certainly limited compared to AGI.

Commercial Machine Learning applications are now deeply embedded in our every-day lives. For example, when you unlock your smartphone with face-recognition, when you speak to Alexa or Google Home to order online or to listen to music, when your bank automatically blocks your credit card for security reasons or when your email client reminds you that you might want to respond to an older message etc., a Supervised Learning algorithm powers the application. More specifically, it performs what is known as classification. In the face-recognition case the algorithm classifies the image as you or not-you. In the virtual assistant case, the algorithm classifies your intent to purchase something online or play a game and so is able to provide you with options in a seemingly conversational manner (a conversational UI).

Another very commonly used capability of machine learning is to measure how similar objects are between them. This is often achieved with Unsupervised Learning algorithms and often with a family of algorithms that perform Clustering. There is an important difference between classification and clustering. The first involves pre-defined types. In contrast, the second assumes that all objects under examination are of the same type and looks to find how similar two or more of them are. Examples of this type you encounter when a social network identifies how similar two users are, in order to suggest people they may know or they may want to connect with within the network, when Netflix and Spotify identify how similar user preferences are so as to suggest relevant movies and music correspondingly, when Amazon recommends products you may be interested in or when you are shown targeted advertisements based on your interests, your searches or your purchasing patterns.

You might be thinking that the breadth of applications is stunning and the applications themselves might seem to you very diverse. This is because they are. So, how can it be that we use the same methods for speech, image, transactions and even people? Going a level deeper, speech for example is the outcome of thought processes, the direct output of real intelligence. Images, on the contrary are just snapshots. And then again, image recognition is something that humans do very well, while recognising a fraudulent transaction would require a certain level of expertise, scrutinising a lot of data and maybe it would be impossible even then. How can the same technology solve for all these different problems?

The latest and perhaps most impressive applications involve Machine Learning systems that are not just able to assort entities but are able to take actions. Example applications are self-driving cars such as the ones of Tesla and Uber, and AI game championing programs such as Google's AlphaGo. Go is a popular board game that is particularly widespread in Korea where it is even taught in school since very young ages. Two players compete on a 19x19 grid and the goal is to place your stones in a way that surrounds larger areas than those surrounded by your opponent. The set of all possible board positions is calculated to be 10^170, a number much larger than the number of atoms in the observable universe (10^80). Go was until recently one of the areas where computers were considered to be very unlikely to beat humans at any point soon. The reason being that it is an extremely difficult game which experts and funs consider to require a lot of creativity and human intuition. This all changed in 2016 when AlphaGo, an AI created by Google's DeepMind won Lee Sedol, a master of the game who ranked at the highest level of skill in the game globally (9 dans). In 2017, AlphaGo moved on to also beat Ke Jien, the #1 master in the world at the time.

This is exactly the state where AI stands today: It can solve a wide array of important vertical problems extremely well. Even more interestingly on the technical level, we have invented families of algorithms that are generic enough to solve for problems that initially at least may not strike us as connected, intuitively. Classification and clustering emerge as a unifying approach or at least important building blocks applicable to all these different problems. We will see why momentarily. On the other hand, we have not come up with an umbrella technology general enough to be able to solve the breadth of problems that humans are able to. ML algorithms are multi-specialists but not generalists, unlike the human mind.

The *Learning* In Machine Learning

Let's take a moment to discuss what exactly this leap means. What makes these new marvelous Machine Learning algorithms, that are able to solve all these problems, stand out compared to the arsenal that we had before, why are they so different.

The reason is that traditionally computers function in a very deterministic and prescribed way. The programmers establish a precise sequence of actions, i.e. an algorithm, and computers execute the algorithm extremely efficiently. The computer's power has been to execute vast repetitive calculations that may involve large numbers in superhuman speed and accuracy. The most interesting point is how computers have been making decisions. The most common decision a computer is asked to take is based on whether two values are exactly equal or one of them is larger. And that's about it: superfast mathematical calculations and arithmetic comparisons (even comparing non-arithmetic parameters such as text all other objects in Object Oriented programming) that basically resolve to exactly equal or not, true for equal or otherwise false. The programmer must predict all possible behaviours and paths that the algorithm can follow, based on this simple equal or not decision mechanism, and prescribe the algorithm's behaviour. If the programmer has not been thorough enough to cover all cases, the program may have "bugs" which may cause it to present unwanted behaviour, become unresponsive, fail, or even crash the entire system. This may sound tedious and boring but software companies have been able to produce a universe of incredible programs with just these two facilities since the decade of 1980.

The real world however is much more noisy, ambiguous and dynamic. Think about the face recognition feature of a smartphone: No scan of your face will be exactly the same to a previous for a multitude of reasons: The angle will always be slightly different, so as the light, your expression and the background behind you. Even if somehow magically you achieved to produce the exact same scan, your face is changing day by day, you may be unshaved, take a haircut and you simply always age. So the pixels in one scan of your face will be greatly different from the pixels of every other one. Despite all this factors that introduce ambiguity, your smartphone is still able to recognise you.

And this is the big difference that Machine Learning has introduced: Computers have gone beyond just been able to say if two objects are strictly equal or not. They can tell if two complex objects are similar and in the case of the face recognition powered phone unlocking, if the two scans are similar enough to belong to the same person with a very high probability to render the feature secure. What we essentially described is the ability to recognise patterns and perform comparisons which involve ambiguity and noise, a very different and much more realistic situation compared to binary equality comparisons.

We will spend the rest of this book explaining exactly how this is done but let's think of what it means on a high level for your smartphone to be able to tell if a scan that it has never seen before belongs to you or your friends. Or to make the point even more evident, you may have seen applications that are able to detect and identify entities, like vehicles in street photos, animals in landscapes etc., without ever having seen before these particular photos. What does this tell us? That the algorithm is able to *learn* to distinguish your face from your friends in the face recognition application or the cars the animals from anything else that is depicted in an image in the entity recognition application. It builds an internal representation that allows it to distinguish between you and a stranger or a car and an animal. It is able to retrieve it and compare it successfully to a new, previously unseen instance of a scan of you or a street or landscape photo respectively.

What Is Knowledge?

At this point you may be wondering: "Really? Do they really *learn*?" Let us unpack how the concept of knowledge can be approached and see how it compares with what Machine Learning algorithms do. Knowledge is the examination and analysis of data concerning a scoped context (subject, process or situation). This analysis may result in a formula that describes a quantity which we are interested in. For example, Eratosthenis calculated the Earth's circumference (at around 250BC with 99.85% accuracy) by using a rod, its shadow and some pre-existing knowledge: basic geometry.

Of course, not everything can be expressed as a single equation. For a more complex phenomenon we may require a system of equations, algorithms or a combination. When such an abstract device, is fed with data about the real phenomenon, a useful representation of this reality is fleshed out. It now embodies our knowledge about the real phenomenon and it is called a *model*.

Deterministic and periodic phenomena are predictable by nature. For example, astronomical positions could be predicted already since ancient Greece (with a device named the Antikithira mechanism). Other phenomena are much more noisy, random and ambiguous. For example weather forecasting is an inherently much less regular and predictable system. In such cases of complex, noisy phenomena, a useful model is a simplified mathematical description of reality, good enough to be able to make practical decisions (e.g. your face scan is a match or not with your real face) or give practical predictions (e.g. tomorrow is going to be a sunny or rainy) in subsequent occasions within the same context.

This is exactly what ML algorithms do. For any given application context, we can pick the ML algorithm we wish to apply, feed it with data and it will learn a model, custom to the problem and able to make decisions about the context. For example, if we choose to implement the finger print unlock feature of a smartphone with a Neural Network algorithm, we feed it with data by applying our finger to the touch id sensor for a number of times and it learns the Neural Network based model of what our fingerprint looks like. Neural Networks work by analysing data, detecting and leveraging patterns in the data, by looking into myriads of interconnected “microscopic” correlations to build models for the entities they are asked to recognise.

This raises the question of how we choose which algorithm to apply. As we discussed some of the most successful Machine Learning algorithms are in principle application agnostic, at least to a certain extent. For example, you can apply artificial Neural Networks to potentially any classification application e.g. speech, image or financial transaction related. What does this imply? That they are not designed to replicate the mechanics of any of the underlying real processes. In all these applications, there is an emerging pattern: the classification, clustering, regression, inference etc. Hence, with artificial Neural Networks for example, we have not built an artificial human brain, rather we have invented a bio-inspired, data-driven algorithm to classify items of potentially any kind, which draws analogies with how the human brain learns. For example, scientists have found that each of the human brain’s areas have the capability to learn different processes (like speech, image recognition etc.) when they rewired sensory input accordingly. The crucial analogy drawn here is that in these experiments the miracle of learning came from the sensory input data interacting with an agnostic learning device (the human brain).

If you are not completely satisfied by this definition for knowledge, stay tuned: We will shortly discuss an alternative one, which is more aligned with a humanist angle. However, bear in mind that the one we discuss currently not only is a scientifically accepted one but has gradually become the dominant perspective, for reasons that will also be explained.

A Ubiquitous Science

So, data is the fuel that propels the power of Machine Learning algorithms. And these algorithms, multi-specialist as they are, can be applied to many domains, e.g. e-commerce, finance, digital security, robotics, medicine, education, communications, virtual assistant technology, law etc.

because they all have applications that boil down to the types that Machine Learning is performing well.

This in turn creates a Ubiquitous Science: Data Science. Data Science is a unifying science because it facilitates a consistent approach to solve for all these different vertical problems. It is the science of determining if the "oil" deposit, i.e. the data, is of practical value, extracting it, distilling it and refining it from the noisy real-world process that generates it. Furthermore, it is about choosing which motors, i.e. which ML algorithm, they will apply to any given situation or even engineering new ones. ML algorithms are generalists but not all apply to any given problem with equal success and each one retains its own characteristics that make it variably applicable to different application scenarios. The principles of Data Science apply horizontally to all application domains, making Data Scientists one of the most in-demand professionals in the global labour market of the time.

If you think closer about it, Dara Science being a unifying science makes total sense: The most fundamental principle of Science, perhaps even the defining one, is the examination of data, the building of hypotheses (a funkier term for models) and then the testing of these hypotheses with new data. As we will see very soon, this is exactly what data scientists do. Hence, Data Science is on a high level the array of techniques for the application of the most fundamental principle of Science.

Finally, the fact that computers can now function successfully in the more ambiguous settings of the real world may result in the line between reality and the cyberspace to start blurring. You may have heard the idea expressed by some scientists as well as technology entrepreneurs that we might be living in a simulation. This may seem far-fetched and rather academic right now, but it might also be how things unfold in the future. We will discuss this scenario in the context of AGI and the Singularity in following chapters.

The Global Data Movement

--- Omitted in this version ---

Is Intelligence In the Data Or In The Mind?

--- Omitted in this version ---

Why ML Took Off Now?

Two other reasons, except the inventions of algorithms that deal with the ambiguity of reality played in to result in Machine Learning skyrocketing recently.

The first one is the tremendous increase in computing power particularly with the commercialisation of the Cloud infrastructure (mainly Amazon AWS and the Google Cloud Platform) but in personal hardware as well. The cloud is the infrastructure, the avenue on which the Machine Learning, data intensive applications can run. It provides with elastic processing power, in simpler terms: dynamically adjustable computing power aggregated and put in use by clusters of many computers.

In turn, two factors played a role in this:

- First the famous Moore's law which dictates that computing power doubles every 18 months to two years. This is a very fast rate. For comparison, for the past few centuries population has doubled every roughly 70 years, per capita consumption every roughly 35 years and scientific progress every roughly 15 years. However, this rate has been constantly happening since the invention of consumer electronics, so you might be wondering why its effect kicked in now. The reason is that doubling a value, then taking the outcome and doubling it again and repeating this multiple times, very soon enters the territory of gargantuan numbers and there is a point at which this becomes a game changer. We are at that point. It is difficult for the human brain to fathom the power of repeated doubling. There is famous myth about the origin of chess that gives perspective: The inventor of chess showed it to the emperor of India, and the latter was so impressed about the brilliant game that asked the inventor to name his reward. The inventor asked that he is given a grain of rice placed at the first square of the chessboard, two at the second, four at the third, eight at the fourth etc. The emperor agreed only to be informed a few days later by his treasurer that to fulfill the ask, they would need all the rice produced in the country for several centuries. With 64 squares on the chess board, 2^63 is an unimaginably large number of rice grains (9 quintillions, i.e. 9 followed by 18 zeros, a number the scale of which human experience is not familiar with). For perspective, it would weigh approximately 280 billion tons (still difficult to fathom).

72P	144P	288P	576P	1E	2E	4E	8E
281T	563T	1P	2P	5P	9P	18P	36P
1T	2T	4T	9T	18T	35T	70T	141T
4B	9B	17B	34B	69B	137B	275B	550B
17M	34M	67M	134M	268M	537M	1B	2B
66K	131K	262K	524K	1M	2M	4M	8M
256	512	1K	2K	4K	8K	16K	33K
1	2	4	8	16	32	64	128

  The point is that mid-way through the chess board, the number becomes extremely big and shortly after, astronomical. It is not intuitive that 1 becomes vast after just thirty doublings or so. In Ernest Hemingway's novel *The Sun Also Rises* two characters discuss:
    - "How did you go bankrupt?"
    - "Two ways. Gradually and then suddenly."

  That's the feeling of the exponential explosion. We don't see it coming and then suddenly reality becomes cataclysmic. The values at the second half of the chess board render the previous ones negligible. This is why the effect of Moore's law is relevant now more than ever before: we may be entering the range where values become important, making the progress up to now negligible, and our intuition capacity "bankrupt".
- Second, big innovators such as Google started experimenting with using multiple commodity cheap machines instead of buying a small number of very expensive super-computers that quickly lose their value as technology improves. This led to a new distributed paradigm for computing (e.g. Hadoop/Map-Reduce, No-SQL databases etc.) which enables programs to run and process data on multiple machines simultaneously so as to deliver service to their massive user bases.

With the number of mobile devices in operation exploding, there is currently a race between the Cloud's computing power and the astronomical flows of data produced by smartphones (which now include almost everyone in western societies), smartwatches and other wearables, the Internet of Things which includes devices such as smart thermostats, smart doorbells etc, virtual assistants, cameras, drones etc. As the mobile devices power increases, engineers are looking to spread Machine Learning across the Cloud and the devices, in order to strike the right balance of processing load between them.

The second reason why ML took off now is the abundance of data, the fuel for ML, which is caused by the ever expanding digitisation of information. Digital content is cheaper to produce and free to reproduce so it has taken over the world and has become the dominant form for multimedia and the press. Books, music, movies and all types of specialised knowledge bases are digitised. The low to no cost distribution of information also fueled a tsunami of user-generated content, through crowd-based projects such as Wikipedia, blogging and social platforms such as Facebook, Twitter and Instagram.

Why ML's Take-Off Is Already A Game-Changer

So, with Machine Learning we have not built human level intelligence but we have created algorithms that can acquire knowledge and make decisions much more complicated and precise than previous algorithms were capable of. In addition, they do so in real-world settings where there is a lot of ambiguity and noise. This has major mid-term implications.

The reason is that in doing so, algorithms can replace humans in many professional situations. Self-driving cars will reduce the need for professional human drivers. Chatbots will reduce the need for sellers, personal assistants, receptionists, customer service agents etc. They will also reduce the need for car and parking estate ownership, since when a human driver is not required, a car can be a shared resource. Robots will reduce the need for builders, warehouse and factory workers. They may also reduce the numbers of human soldiers or even officers policing the streets. Industrial robots are out there for quite a long now but their modern days ancestors are able to function in real world environments, immune to the natural noise and ambiguity. No longer does a robot which works on a transport belt or assembly line needs an object to be at a precise position in order to be able to pick it up. Current robots can detect items of variable sizes and positions and adjust their angle of attack and grip as needed. A famous example is Atlas, the parkour capable robot by Boston Dynamics.

Traditionally coveted, high earning jobs are not immune either: AI driven medical networks may start outperforming human diagnosticians. IBM for example is working on a healthcare version of Watson, the AI program that beat the top players of Jeopardy in 2011 already. If you think about it, examining the symptoms in order to produce a diagnose is a classification task. Surgeons, lab workers or other practitioners may be threatened by robots. AI powered lawyers will be able to read contracts and assess the risks for their clients, just another classification task.

There are two very visible implications, if this unfolds accordingly:

First, productivity may skyrocket as the software and robots will be much cheaper than the human workers whom they will replace, they will work much faster and more accurately. In addition, they will be much faster to train (human doctors often study for more than a decade after school) and, lest you forget, they will also be extremely loyal. Except for the AI dooms day scenario materialises of course....

Second, if scores of people are bound to lose their jobs and, even worse, their place in the labour market on a more permanent basis, the entire economic system will undergo dramatic changes and the very fabric of our society's organisation may be disrupted. We will dedicate an entire chapter on the longer term and more structural implications of an upcoming Intelligence revolution. At this point, let us try to investigate how likely it is.

How Today's ML May Lead To Human-Level Intelligence

--- Omitted in this version ---

The Hardware Of Intelligence

--- Omitted in this version ---

An Alternative Path To Human-Level Intelligence: Whole-Brain Emulation

--- Omitted in this version ---

The Historical And Economic Signs That An AI Revolution Might Be On The Horizon

Are we due to face an AI revolution soon? Let's first try to understand what a "revolution" means in the context of human history, identify characteristics and then turn to what an intelligence revolution particularly may look like and what would be some reshaping consequences.

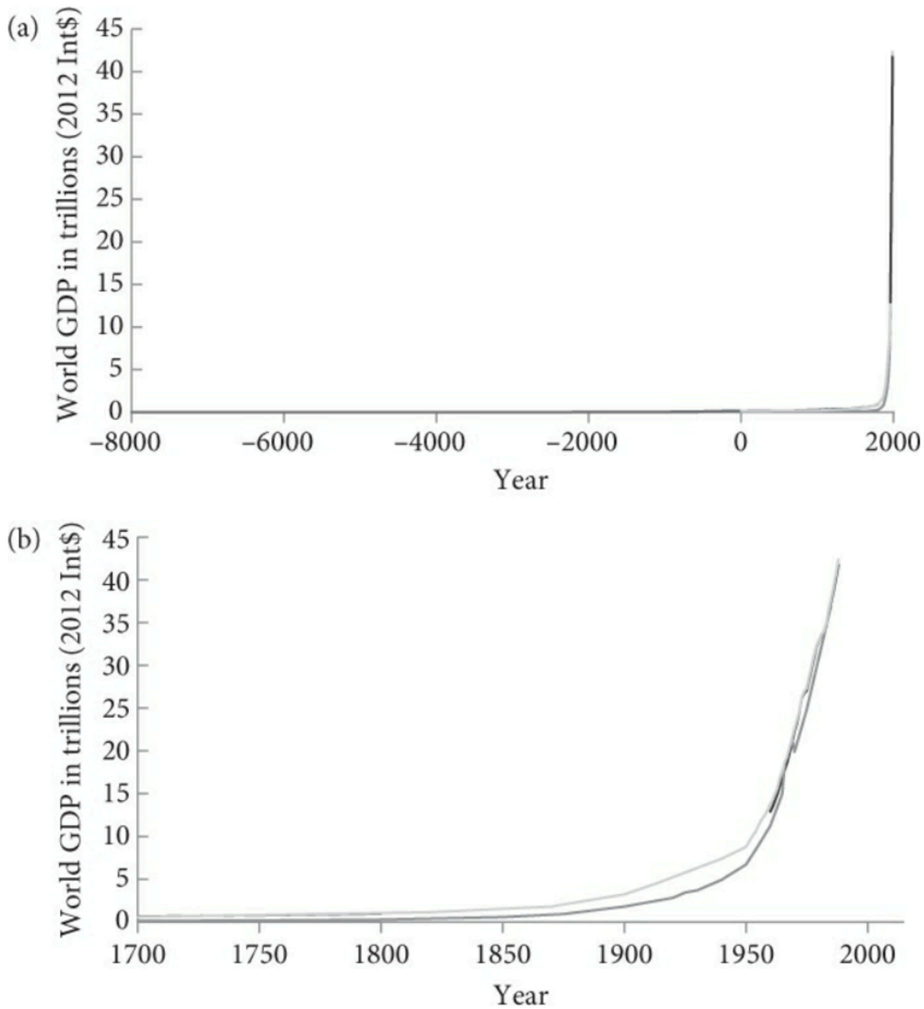
Graph X is a very important graph, it summarises humanity's economic history. If you look at the big history sub-graph, around 1900 there is a turning point in human economic history. It is when the industrial revolution started. Here you can observe the exponential explosion concept we discussed earlier. Looking from the industrial revolution turning point, all previous developments look flat and insignificant.

If you examine the second subgraph that zooms in to the latest period, you will also see that the effect of the industrial revolution took some time to kick in, as the technologies matured and the exponential growth was building up. As part of the industrial revolution, the invention of the steam engine at around 1770 is probably economically the most important event in economic history yet.

Economic historians model the evolution of human economy by proxy of the human collective product. The historical human product fits the model of exponential growth with different exponents for the different historical periods of hunting, farming and industrialisation. In other words, the product trend can be modeled as a series of exponential functions with a step function that increases the exponent dramatically when from hunting we transitioned to farming and from farming to industrialisation. The first evidence of hunting activity goes back to almost two million years ago. In the hunting period, the product of human activity doubled every 230,000 years. During the epoch of hunting, humans were nomadic and egalitarian.

At around 5,000 BC there is a large-scale transition from hunting to agriculture. During this period the human product doubles every roughly 900 years. During the period of farming, humans settled, shaping villages and cities, shaping the natural environment and domesticated animals.

With the industrial revolution, from roughly 1800 to 1900 the human product doubles every roughly 58 years. From 1900 to 2000 the human product doubles every roughly 15 years and it is projected to grow even faster from 2020 onwards. The industrial revolution introduces technology and largely shifts us from feudalism to capitalism. Not coincidentally, the global population also increased its growth rate at these very same transitional historical points.



Graphs from “Superintelligence: Paths, Dangers, Strategies”, by Nick Bostrom, 2014

Hence, the two major consequences we identified as possible outcomes of the Machine Learning take-off are both present and prominent characteristics of past revolutions: First, productivity skyrockets compared to the previous epoch. Second, the economic system gets disrupted and society is reshaped drastically.

Have we entered the “waiting room” of a new historical period, the period of the Intelligence revolution, and we are only due to experience the effects, just like humanity had to wait a few decades to realise the effects of the industrial revolution?

There is research suggesting, by extrapolating the model of the series of exponentials, that a new transition will happen within the twenty-first century, which will cause the economy to double every few weeks, instead of every few years currently. Of course, this is just a model and we have not seen enough such transitions to be able tell if the model’s extrapolation is trustworthy.

If the scenario of the global economy doubling every few weeks materialises at any point, within the century or beyond, it is difficult to fathom the new reality. It would cause impossible to predict cataclysmic transformations to the human civilisation and make the industrial revolution period to look flat and insignificant. This outcome goes by the term: Singularity. The term originates in Mathematics and Physics. In mathematics, a singularity is a point where a quantity is not defined. For example, the function $1/x$ has a singularity at $x=0$. Somewhat more complicated, in General Relativity a gravitational singularity is an area where the density of the matter becomes infinite, causing the laws of Physics to collapse. This happens at the center of black holes and it happened at the birth of the Universe during the Big Bang. By analogy in the event of the computational singularity, the human affairs as we understand them will not continue.

The Economics Of AI

--- Omitted in this version ---