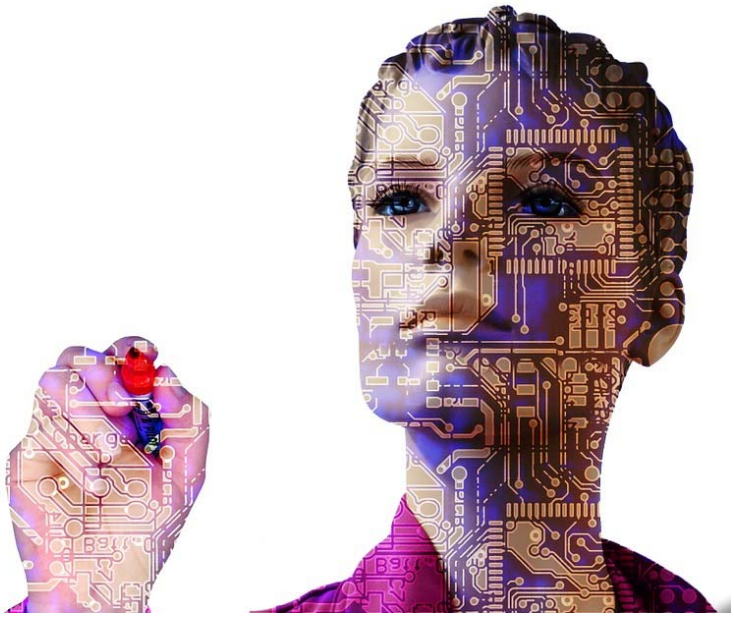

The illusions of AI in 2016: An honest comparison with the 80's

I am studying and teaching AI by the 80's and, although I am still an AI enthusiast, I feel compelled to express some doubts. I am aware that I could be wrong, so please do not confuse my criticism with defeatism.

We are currently submersed by messages that speak about "machine learning", "big data analysis", "self-driving cars" or "intelligent apps" for smartphones that can recognize our mood.

It seems that AI has made strides from the 80's when AI failed the overly optimistic expectations. Is everything true? Or should we try to understand better what is behind these slogans and promises? I believe the second option is correct: It is a learned lesson from the 80's.

The world has deeply changed from the 80's and the fields of applicability of AI has changed and expanded. The Internet produced large amount of data that, in the 80's, were almost inconceivable. The technology, particularly the hardware of computers, has made strides and thus the instruments needed to execute AI algorithms are much more powerful. The question is: "what improvements have been done in AI in the last thirty years?".

In the 80's, the Expert Systems were the AI consolidated technology while the Neural Networks were the state of the art, more exotic because brain-inspired and more promising due to their capability to learn, not present in Case Based Reasoning Systems. The Error Back Propagation Algorithm [1] helped a lot the take-off of the Neural Networks field, while other algorithms, like the Adaptive Resonance Theory [2], solved the plasticity versus stability problem, and the hierarchical NeoCognitron [3] became the precursor of the current Convolutional Neural Networks [4].

The big error in that period has been the lack of a serious analysis of the potentiality of the new instruments and the result has been the growth of overly optimistic expectations in the common people. In such a context, despite the fact that AI had many valid argumentations, the failure was unavoidable. In the following thirty years, the field of Neural Networks has grown, differentiating in multiple disciplines with different targets from neuroscience to engineering. The engineering applications, like pattern recognition and data analysis, are still classified as Artificial Intelligence. There have been some improvements in algorithms for pattern recognition: Convolutional Neural Networks have been an important development in artificial vision while Echo State Networks have demonstrated usefulness in other tasks. The deep learning, with hierarchical structures, has been made possible by the use of high performance GPUs. But please note that Deep Learning is not a new concept: the first functional Deep Learning networks with many layers were published by Alexey Grigorevich Ivakhnenko and V. G. Lapa in 1965 [5]. The third Generation Neural Networks (Spiking Neural Networks) [6] are a promise for the future but are not still useful in engineering applications. They are not efficient on Von Neumann computers and the current state of the art of spiking neural hardware (IBM TrueNorth) [7] still misses some fundamental capabilities like on-chip learning and self-organizing behavior: currently this chip needs to be programmed using an off-line learning process. I am expecting big improvements in this field and probably AI will have a paradigm shift thanks to hardware research. Although the trend in the design of neuromorphic chips is using IF (Integrate and Fire) neurons that mimic biological neural networks, it is not yet demonstrated that this model is the best choice for designing artificial brains on silicon: as example Switching Neural Networks [8] could be a good candidate because they are universal approximators and are optimized for digital implementations. It is interesting to note that in the 80s/90s there were many neuromorphic (not spiking) chips commercially available (Micro Devices MD-1220, NeuraLogix NLX-420, Philips Lneuro-1, Nestor/Intel NI1000, Silicon Recognition ZISC-78 and many others) [9], but for my best knowledge, only one survived on the market: the ZISC-78 evolved to the current CM1K (General Vision) with the same RBF architecture (on-chip learning capability) and much more neurons. I guess that the reason is that the company producing this chip has been focusing on the development of practical applications (like image recognition) rather than proposing improbable sci-fi solutions. A new commercial chip (Emoshape) synthetizes emotions (with an approach that is more complete than a simple competitive model) and seems to be interesting for the development of AI based robotics. I have not yet enough technical details to build a technical opinion on the chip: currently, I can say that the approach could be not compliant with a distributed representation of emotions (like in a connectionist model). The interaction with the human needs emotions, but my opinion is that emotions make sense in a complex network of neuro-cognitive networks, where the emotions could interact and influence flows and behaviors. Emotions in a simple toy or a smartphone are just matter for marketing slogans.

Self-driving cars still use an ensemble of deep neural networks and case based reasoning engines: they can perform correctly only in low traffic roads with favorable weather conditions. They are still far from the flexibility, adaptability and the full features learning capability of a human driver. Furthermore, driving a car is not only matter of pattern recognition and fast reactions but it is also matter of ethics. A driver should decide a priority about the safety of passengers and pedestrians and he should take this decision on the base of many factors. This is a hard task for humans and it is almost unconceivable for the current state of the art of AI. On the other side, the positive fact is that computers on a self-driving car, differently from humans, are not distracted by OS processes that send movies to YouTube or messages with WhatsApp. I hope that this feature will be never added to the self-driving cars software!

Some start-up offers “intelligent apps” for the smartphones: as example these apps can understand emotions of people that look to a video on the smartphone and change the video consequently. Useful? My opinion is not important: the success will be determined by the people (perhaps through the creation of a false need, operated by a persuasive advertising campaign). Technically, I can only state that emotions from face expressions can be simply detected with pattern recognition algorithms no more sophisticated than the algorithms existing in the 80’s: the power of the current processors makes the difference.

The discipline called “machine learning” is living a golden period. It mainly uses neural networks and statistical algorithms to process and analyze big data coming from the internet, with marketing goals. Other targets are also pattern recognition, text and image analysis. Here there are two new elements with respect to the 80’s: the large amount of data coming from the internet and the power of the new microprocessors. The algorithms are not actually so different except for the fact they are “pre-packed” in software frameworks that do almost everything: they hide the neural networks algorithms and expose more friendly APIs. It sounds good but, unfortunately, Machine Learning experts are almost always required to have familiarity with these frameworks more than having a real culture on neural networks, statistics and pattern recognition algorithms: are we moving towards this kind of pseudo-scientific and not self-contained culture?

Despite algorithmic innovation in the software, the great success of the discipline is mostly due to the following three factors:

- Availability of big data

- Powerful processors (in particular the GPUs for deep learning)

- Demand of data analysis for marketing

The giants of informatics like Google try to acquire any existing company working in this field because it is the new business. And it is the truth! Google can’t be wrong in this. Anyway, technically, we are speaking of logistic regression, linear regression and well known statistical algorithms. What is more intelligent in this AI than in the 80’s?

The word “intelligence” derives from the Latin verb ability to think “intelligere”. Although “intelligence” has been defined in many different ways, the common perception of the word comprises all the cognitive capabilities of the humans. The current AI algorithms are quite far to implement all these capabilities: a GPU that executes, with Teraflops power, a deep neural network learning task, does not have metacognition and even cognition. Metacognition is a higher order thinking. It is the ensemble of processes that we use to plan, monitor and measure our understanding and learning performance. Metacognition is the awareness of ourselves as thinkers and learners. The current models of metacognition applicable to AI systems have a narrow focus, because they do not address comprehensively the elements of metacognition.

The processing speed (also if determined by high parallelism), the number of neurons or layers in a deep neural network are not a measure of intelligence.

In a period where innovations are often just a “rename” of past technologies, why not change the name of AI with one more appropriate? If not, what name will we assign to the science that will produce systems capable of cognition, metacognition, intuition and creativity? AI^2? AI^3?

There will be a future with machines having the same intelligence of humans (and it may be a danger to mankind), but this future is not tomorrow!

We know many parts of the human brain but our knowledge is sectorial. If the human brain is an orchestra, we can now describe how the violin behaves, however we are still in the dark as regards the behavior of the whole orchestra. The Blue Brain Project aims at creating a digital reconstruction of the brain by reverse-engineering mammalian brain circuitry. It is an important initiative, but I believe that the knowledge acquired by this research will be more interesting for neuroscience and less for AI: complex simulations of networks based on biological neurons will not produce behavioral models easily replicable on silicon.

Some startups promise "conscious machines" available on the market in few years and someone claims to have designed a "self-aware processor", while the debate between scientists and philosophers continues to produce disappointing theories on the essence of consciousness.

With the experience of the 80's, we should make some reflections. There is a new big market of lies and delirious claims, fertile ground for ephemeral business and, probably, waste of public money.

Science and research, by a more critical spirit, will avoid jeopardizing their credibility.

Luca Marchese *luca.marchese@synaptics.org*

Bibliography

[1] Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". Nature. 323 (6088): 533–536. doi:10.1038/323533a0.Art

[2] Carpenter, G.A. & Grossberg, S. (2003), Adaptive Resonance Theory, In Michael A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, Second Edition (pp. 87-90). Cambridge, MA: MIT Press

[3] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36(4): 93-202, 1980.CNN

[4] "Convolutional Neural Networks (LeNet) – DeepLearning 0.1 documentation". DeepLearning 0.1. LISA Lab. Retrieved 31 August 2013.

[5] Ivakhnenko, Alexey (1965). Cybernetic Predicting Devices. Kiev: Naukova Dumka.

[6] Maas, Wolfgang (1996). "Networks of Spiking Neurons: The Third Generation of Neural Network Models". doi:10.1016/S0893-6080(97)00011-7.

[7] Merolla, P. A.; Arthur, J. V.; Alvarez-Icaza, R.; Cassidy, A. S.; Sawada, J.; Akopyan, F.; Jackson, B. L.; Imam, N.; Guo, C.; Nakamura, Y.; Brezzo, B.; Vo, I.; Esser, S. K.; Appuswamy, R.; Taba, B.; Amir, A.; Flickner, M. D.; Risk, W. P.; Manohar, R.; Modha, D. S. (2014). "A million spiking-neuron integrated circuit with a scalable communication network and interface". Science. 345 (6197): 668. doi:10.1126/science.1254642.

[8] Muselli, Marco (2006). "Switching Neural Networks: A new connectionist model for classification". WIRN 2005 and NAIS 2005, Lecture Notes on Computer Science. 3931: 23–30.

[9] Fernando Morgado Dias , Ana Antunes, Alexandre Manuel Mota – "COMMERCIAL HARDWARE FOR ARTIFICIAL NEURAL NETWORKS: A SURVEY " - Email:{fmdias,aantunes}@est.ips.pt, Email: alex@det.ua.pt.