The Artificial Intelligence Revolution: Reality or Myth?

Dr. M. Vidyasagar

SERB National Science Chair and
Distinguished Professor, IIT Hyderabad
KID: 20200205

At present, almost everyone has heard phrases such as Artificial Intelligence (AI), Machine Learning (ML), Data Science, and the like. Courses on these topics abound, both on the Internet and in various types of institutions. Various success stories of AI in image recognition, speech recognition, drug discovery, medical diagnosis, etc. can be found in popular media almost every day. Leading academic institutes such as IIT Hyderabad have introduced specialized programs in AI at all three levels: Bachelor's, Master's and Doctoral. If the experience of our graduates (and those from other top institutions around the world) is any indication, industry is ready to snap up the graduates of these programs at highly attractive salaries.

And yet one can ask: Is the AI revolution (as I choose to call it) a myth or a reality? Have there been precursors to the present wave of interest in AI? If so how did the earlier waves play out?

The phrase "Artificial Intelligence" was coined by the Stanford Professor John McCarthy in 1955. The reader may be surprised to know that there have been at least three such "hype cycles" regarding AI since that time. The first hype cycle centered around what was called the "perceptron" which was invented in 1962 by Frank Rosenblatt. Perceptrons were claimed to match the performance of humans at checkers (or draughts as they are called in some countries). But the hype ended in 1969 with the publication of the book "Perceptrons" by Minsky and Papert in 1969. The book showed very clearly that perceptrons were incapable of solving some very elementary problems, and sent the book into hiatus for nearly twenty years. During the 1970s, an entirely different approach was initiated, based on using AI to mimic human reasoning. These computer programs were called "expert systems" and were meant to enable novice humans to perform at the level of expert humans, by capturing the knowledge of the latter in a set of "rules." Expert systems were quite good for solving highly structured problems such as diagnosing faults in a radar for example. However, they stumbled when confronted on real-world problems that did not strictly follow the rules of logic, and instead required intuition and judgement. Thus rule-based expert systems never became anything more than niche solutions to specific problems.

The next hype cycle started in 1986 with the publication of a three-part book titled "Parallel Distributed Processing" by Rumelhart, McClelland and Hinton. Rather ironically, this book revived the perceptron, but in a different architecture called Multi-Layer Perceptron Networks (MLPNs). This class of networks are also known as neural networks. MLPNs could solve more complex problems than single perceptrons. More to the point, mathematical theories of "learning and generalization" were developed to explain *why* MLPNs worked so well. I too have written two books on this topic, one in 1997 and another in 2003. Unlike earlier hype cycles, this one did not collapse. Rather, after the initial excitement, the research area went into a quiescent period, awaiting further development.

Continued...

The current hype cycle in AI began in the current decade, and much of it is based on "deep learning." Specifically, while the MLPNs of the 1990s consisted of a few dozen neurons at three or four layers, current networks consist of hundreds of layers and millions of neurons. The advances in deep learning can be attributed to three factors, in decreasing order of importance:

1. Massive increase in computing power, exemplified by GPUs (Graphical Processing Units) and TPUs (Tensor Processing Units).

2. Availability of enormous amount of data, to train neural networks.

3. Invention of new algorithms.

Notice that I put the invention of new algorithms at the bottom. I believe that the availability of computing power and of data contributed much more to the recent advances in AI than the invention of new algorithms, though the last item is also important.

So where do we stand today? Will the current hype cycle survive like the invention of MLPNs, or collapse like the two preceding cycles? My own belief is that it is too early to say. The availability of cheap and powerful computation tempts the user to build overly elaborate networks to solve the problem at hand. As of now, there is ample empirical evidence, and some theoretical evidence, to show that when overly elaborate network architectures are used to solve AI problems, the performance of the resulting networks is often fragile (changes drastically with minor changes in the training data), easily fooled by imperceptible changes in the input data, and other such shortcomings. The theory of "deep learning" lags the practical application at the moment. This is in contrast with the MLPN era when the theory lagged only a few years behind the practice. It does not help that the complexity of current AI systems makes theoretical analysis very challenging. Nevertheless, if the collapse of the first two hype cycles and the survival of the third hype cycle has shown us anything, it is that discipline will survive when it has a solid mathematical foundation. Thus it is imperative for the research community to continue its investigation into the mathematical foundations of deep learning. This is my personal area of research at the moment.