

Is Deep Learning Approaching a Wall?

As with all major technologies in their early phases, deep learning must overcome a number of serious challenges



MIT IDE

[Follow](#)

Sep 2, 2018 · 6 min read



By Irving Wladawsky-Berger

Deep learning and related machine learning advances have played a central role in AI's recent achievements, giving computers the ability to be trained by ingesting and analyzing large amounts of data instead of being explicitly programmed. In just the past two years, Google's deep-learning-based AlphaGo defeated the world's top Go players, surprising most AI experts who thought that it would take another 5 to 10 years to achieve such a milestone. Similarly, when Google switched to its new deep learning AI system in late 2016, it achieved an overnight improvement in the quality of its machine translations roughly equal to the total gains that the previous program had accrued over its 10 year lifetime.

As is typically the case with major technology achievements—e.g. the dot-com bubble—deep learning has quickly climbed to the top of Gartner's hype cycle, where all the excitement and publicity accompanying new, promising technologies often leads to *inflated*

expectations, followed by *disillusionment* if the technology fails to deliver. AI may be particularly prone to such hype cycles, as the notion of machines achieving or surpassing human levels of intelligence leads to feelings of wonder as well as fear. Over the past several decades, AI has gone through a few such hype cycles, including the so-called AI winter in the 1980s that nearly killed the field.

In a recent article, Deep Learning: A Critical Appraisal, author and NYU professor Gary Marcus offers a serious assessment of deep learning. He argues that, despite its considerable achievements over the past five years, deep learning may well be approaching a wall, an opinion apparently shared by University of Toronto professor Geoffrey Hinton, the so-called Godfather of Deep Learning.

Deep learning is a powerful statistical technique for classifying patterns using large training data sets and multi-layer AI neural networks. It's essentially a method for machines to learn from data that's loosely modeled on the way a biological brain learns to solve problems. Each artificial neural unit is connected to many other such units, and the links can be statistically strengthened or decreased based on the data used to train the system. Each successive layer in a multi-layer network uses the output from the previous layer as input.

“In principle, given infinite data, deep learning systems are powerful enough to represent any finite deterministic ‘mapping’ between any given set of inputs and a set of corresponding outputs, though in practice whether they can learn such a mapping depends on many factors...” writes Marcus. “The technique excels at solving closed-end classification problems, in which a wide range of potential signals must be mapped onto a limited number of categories, given that there is enough data available and the test set closely resembles the training set. But deviations from these assumptions can cause problems; deep learning is just a statistical technique, and all statistical techniques suffer with deviation from their assumptions... In practice, results with large data sets are often quite good, on a wide range of potential mappings.”

As with all major technologies in their early phases, deep learning must overcome a number of serious challenges.

Marcus' [article](#) discusses ten such challenges faced by current deep learning systems. I'll focus my discussions on four in particular.

Deep learning is data hungry

The data requirements for deep learning are [substantially different](#) from those of other analytic methods in a number of dimensions. The performance of traditional analytics tends to plateau as the data set size increases. However, the performance of properly trained deep learning techniques will significantly improve as the data sets get larger. Deep learning methods are particularly valuable in extracting patterns from complex, unstructured data, including audio, speech, images and video. To do so, they require thousands of data records for models to become good at classification tasks and millions for them to perform at the level of humans.

“Human beings can learn abstract relationships in a few trials,...” notes Marcus. “Deep learning currently lacks a mechanism for learning abstractions through explicit, verbal definition, and works best when there are thousands, millions or even billions of training examples.” When learning through explicit definition, “you rely not on hundreds or thousands or millions of training examples, but on a capacity to represent abstract relationships between algebra-like variables. Humans can learn such abstractions, both through explicit definition and more implicit means. Indeed even seven-month old infants can do so, acquiring learned abstract language-like rules from a small number of unlabeled examples, in just two minutes.”

At a recent [AI conference](#), MIT brain and cognitive sciences professor [Josh Tenenbaum](#) explained [his views](#) on the difference between our present state of AI and the long-term quest for human levels of intelligence. Human-level intelligence requires the ability to go beyond data and machine learning algorithms. Humans are able to build models of the world as they perceive it, including practical, everyday [common sense](#) knowledge, and then use these models to explain their actions and decisions. According to Tenenbaum, three month old babies have a more commonsense understanding of the world around them than any AI application ever built. An AI application starts with a blank slate before learning from patterns in the data it analyzes, while babies start off with a genetic head start and a brain structure that allows them to learn much more than data and patterns.

Research efforts like MIT's Human Dynamics Lab and the Allen Institute, and startups like Kyndi are trying to get around the limitations of deep learning by emulating human common-sense reasoning and/or complementing the statistical oriented AI methods with logic-based programming tools. Such research efforts are still in their very early stages.

Deep learning is actually quite shallow

The *deep* in deep learning refers to its highly sophisticated, multi-layered statistical properties. But, while capable of some amazing results, in its present incarnations deep learning is actually quite shallow and fragile. “[T]he patterns extracted by deep learning are more superficial than they initially appear.”

Our present AI applications do just one thing quite well by being trained with lots of data and deep learning algorithms. Each application must be separately trained with its own data sets, even for use cases that are similar to previous ones. There is, so far, no good way to carry training from one set of circumstances to another.

AI does best with applications and test sets that closely resemble those used in the training set, but it does much less well when attempting to generalize or extrapolate beyond its training data sets.

Deep learning is not sufficiently transparent

Why was a certain decision reached? Another important challenge with deep learning is its opacity and *black box* nature. It's quite difficult to explain in human terms the results of complex deep learning applications. Typical deep learning systems have huge numbers of parameters within their complex neural networks. It's very hard to assess the contributions of individual nodes to a decision in terms that a human will understand.

“The transparency issue, as yet unsolved, is a potential liability when using deep learning for problem domains like financial trades or medical diagnosis, in which human users might like to understand how a given system made a given decision... such opacity can also lead to serious issues of bias.”

Deep learning is difficult to engineer

Another major set of challenges is the engineering risks inherent in any complex, bleeding-edge IT system, especially when used in high-stake applications, e.g. medicine, cars, airplanes, finance, and government. While these risks apply to the growing complexity of AI systems in general, they could be particularly problematic with deep learning given their statistical nature, opacity, and difficulty distinguishing causation from correlation.

We must also ensure that our complex AI systems do what we want them to do and behave the way we want them to behave, a particularly tough problem with deep learning algorithms that are trained with and learn from data without being explicitly programmed.

A number of initiatives have been organized to address these and other major AI and deep learning challenges, including Stanford's One Hundred Year Study of AI, and MIT's Quest for Intelligence. Hopefully — as has been the case with previous powerful technologies — such efforts will help ensure that these challenges are properly addressed, and that our increasingly capable AI systems will have a major beneficial impact on the economy, society and our personal lives.

. . .

Originally published at blog.irvingwb.com on September 3, 2018.

