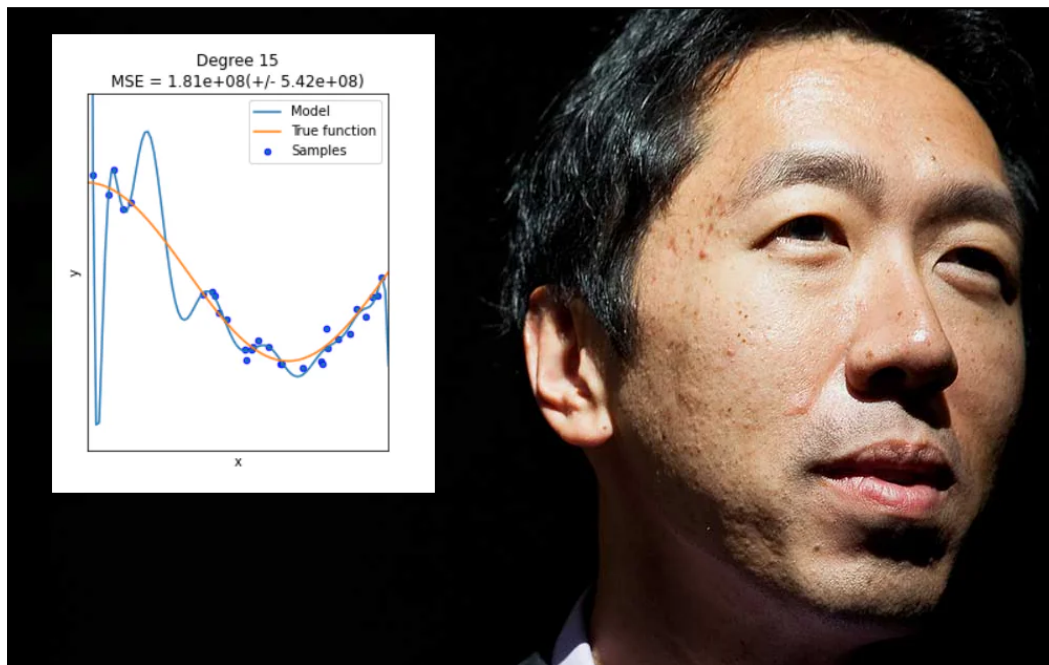


Andrew Ng Criticizes the Culture of Overfitting in Machine Learning

Originally published at <https://www.unite.ai/andrew-ng-criticizes-the-culture-of-overfitting-in-machine-learning/>



Andrew Ng, one of the most influential voices in machine learning over the last decade, is currently voicing concerns about the extent to which the sector emphasizes innovations in model architecture over data – and specifically, the extent to which it allows ‘overfitted’ results to be depicted as generalized solutions or advances.

These are sweeping criticisms of current machine learning culture, emanating from one of its highest authorities, and have implications for confidence in a sector beset by fears over a [third collapse](#) of business confidence in AI development in a space of sixty years.

Ng, a professor at Stanford university, is also one of the founders of deeplearning.ai, and in March published a [missive](#) on the organization’s site that distilled a [recent speech](#) of his down to a couple of core recommendations:

Firstly, that the research community should stop complaining that data cleaning represents 80% of the challenges in machine learning, and get on with the job of developing robust MLOps methodologies and practices.

Secondly, that it should move away from the ‘easy wins’ that can be obtained by over-fitting data to a machine learning model, so that it performs well on that model but fails to generalize or to produce a widely deployable model.

Accepting The Challenge Of Data Architecture And Curation

‘My view’, Ng wrote. ‘is that if 80 percent of our work is data preparation, then ensuring data quality is the important work of a machine learning team.’

He continued:

'Rather than counting on engineers to chance upon the best way to improve a dataset, I hope we can develop MLOps tools that help make building AI systems, including building high-quality datasets, more repeatable and systematic.'

'MLOps is a nascent field, and different people define it differently. But I think the most important organizing principle of MLOps teams and tools should be to ensure the consistent and high-quality flow of data throughout all stages of a project. This will help many projects go more smoothly.'

Speaking on Zoom at a live streamed [Q&A session](#) at the end of April, Ng addressed the applicability shortfall in machine learning analysis systems for radiology:

"It turns out that when we collect data from Stanford Hospital, then we train and test on data from the same hospital, indeed, we can publish papers showing [the algorithms] are comparable to human radiologists in spotting certain conditions.

"...[When] you take that same model, that same AI system, to an older hospital down the street, with an older machine, and the technician uses a slightly different imaging protocol, that data drifts to cause the performance of AI system to degrade significantly. In contrast, any human radiologist can walk down the street to the older hospital and do just fine."

Under-specification Is Not A Solution

Overfitting occurs when a machine learning model is specifically designed to accommodate the eccentricities of a particular dataset (or of the way that the data is formatted). This can involve, for instance, specifying weights that will produce good results from that dataset, but will not 'generalize' on other data.

In many cases, such parameters are defined on 'non-data' aspects of the training set, such as the specific resolution of the gathered information, or other idiosyncrasies that are not guaranteed to re-occur across other subsequent datasets.

Though it would be nice, overfitting is not a problem that can be solved by blindly widening the scope or flexibility of data architecture or model design, when what is actually needed are widely applicable and highly salient features that will perform well across a range of data environments – a thornier challenge.

In general, this type of 'under-specification' only leads to the very problems that Ng has lately outlined, where a machine learning model fails on unseen data. The difference in this case is that the model is failing not because the data or data formatting is different from the overfitted original training set, but because the model is too flexible rather than too brittle.

Late in 2020 the [paper](#) *Underspecification Presents Challenges for Credibility in Modern Machine Learning* leveled intense criticism against this practice, and bore the names of no less than forty machine learning researchers and scientists from Google and MIT, among other institutions.

The paper criticizes 'shortcut learning', and observes the way that underspecified models can take off at wild tangents based on the random seed point at which the model training begins. The contributors observe:

'We have seen that underspecification is ubiquitous in practical machine learning pipelines across many domains. Indeed, thanks to underspecification, substantively important aspects of the decisions are determined by arbitrary choices such as the random seed used for parameter initialization.'

Economic Ramifications Of Changing The Culture

Despite his scholarly credentials, Ng is no airy academic, but has deep and high-level industry experience as the co-founder of Google Brain and Coursera, as former chief scientist for Big Data and AI at Baidu, and as the [founder](#) of Landing AI, which administrates \$175 million USD for new startups in the sector.

When he says “All of AI, not just healthcare, has a proof-of-concept-to-production gap”, it’s intended as a wake-up call to a sector whose current level of hype and spotted history has increasingly characterized it as an uncertain long-term business investment, [beset](#) by problems of definition and scope.

Nonetheless, proprietary machine learning systems that operate well in-situ and fail in other environments represent the kind of market capture that could reward industry investment. Presenting the ‘overfitting problem’ in the context of an occupational hazard offers a disingenuous way to [monetize](#) corporate investment in open source research, and to produce (effectively) proprietary systems where replication by competitors is possible, but problematic.

Whether or not this approach would work in the long term depends on the extent to which real breakthroughs in machine learning continue to require [ever-greater levels of investment](#), and whether or not all productive initiatives will inevitably migrate to FAANG to some extent, due to the colossal resources necessary for hosting and operations.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More