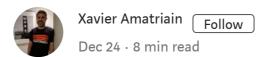
The year in AI/ML advances: 2018 roundup



It has become a sort of tradition for me to try to summarize ML advances at this time of the year (see here for my Quora answer last year, for example). As always, this summary will necessarily be biased by my own interests and focus, but I have tried to keep it as broad as possible. Note that what follows is a blog post version of my Quora answer here.

If I had to summarize the main highlights of machine learning advances in 2018 in a few headlines, these are the ones that I would probably come up:

- · AI hype and fear mongering cools down
- More focus on concrete issues like fairness, interpretability, or causality
- Deep learning is here to stay and is useful in practice for more than image classification (particularly for NLP)
- The battle on the AI frameworks front is heating up, and if you
 want to be someone you better publish a few frameworks of your
 own

Let's look at all of this in some more detail.

If 2017 was probably the cusp of fear mongering and AI hype, 2018 seems to have been the year where we have started to all cool down a bit. While it is true that some figures have continued to push their message of AI fear, they have probably been too busy with other issues to make of this an important point of their agenda . At the same time, it seems like the press and others have come at peace with the idea that while self-driving cars and similar technologies are coming our way, they won't happen tomorrow. That being said, there are still voices defending the bad idea that we should regulate AI instead of focusing on regulating its outcomes.

It is good to see that this year though, the focus seems to have shifted to more concrete issues that can be addressed. For example, there has been a lot of talk around **fairness** and there are not only several conferences on the topic (see <u>FATML</u> or <u>ACM FAT</u>) even some online courses like <u>this one</u> by Google.

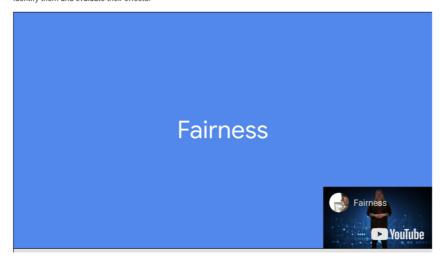
Fairness

☆☆☆☆



Evaluating a machine learning model responsibly requires doing more than just calculating loss metrics. Before putting a model into production, it's critical to audit training data and evaluate predictions for bias.

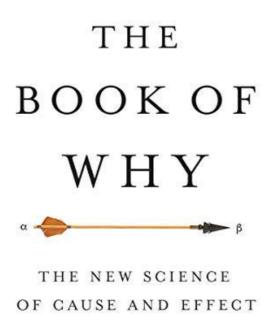
This module looks at different types of human biases that can manifest in training data. It then provides strategies to identify them and evaluate their effects.



Fairness online course by Google

Along these lines, other issues that have been greatly discussed this year include interpretability, explanations, and causality. Starting with the latter, causality seems to have made it back to the spotlight mostly because of the publication of Judea Pearl's "The Book of Why". Not only the author decided to write his first "generally accessible" book, but he also took to Twitter to popularize discussions around causality. In fact, even the popular press has written about this as being a "challenge" to existing AI approaches (see this article in The Atlantic, for example). Actually, even the best paper award at the ACM Recsys conference went to a paper that addressed the issue of how to include causality in embeddings (see "Causal Embeddings for Recommendations"). That being said, many other authors have argued that causality is somewhat of a theoretical distraction, and we should focus again on more concrete issues like interpretability or explanations. Speaking of explanations, one of the highlights in this area might be the publication of the paper and code for Anchor, a follow up to the well-known LIME model by the same authors.

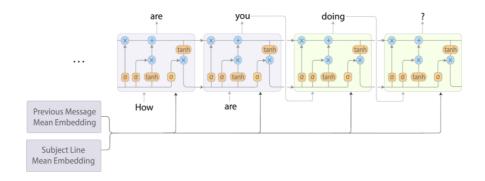
JUDEA PEARL WINNER OF THE TURING AWARD AND DANA MACKENZIE



Instant classic by Judea Pearl

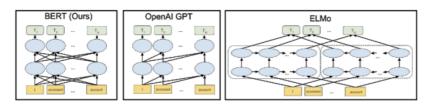
While there are still questions about the Deep Learning as the most general AI paradigm (count me in with those raising questions), while we continue to skim over the nth iteration of the discussion about this between Yann LeCun and Gary Marcus, it is clear that Deep Learning is not only here to stay, but it is still far from having reached a plateau in terms of what it can deliver. More concretely, during this year Deep Learning approaches have shown unprecedented success in fields different from Vision, ranging from Language to Healthcare.

In fact, it is probably in the area of **NLP**, where we have seen the most interesting advances this year. If I had to choose the most impressive AI applications of the year, both of them would be NLP (and both come from Google). The first one is <u>Google's super useful smart compose</u>, and the second one is their <u>Duplex</u> dialog system.



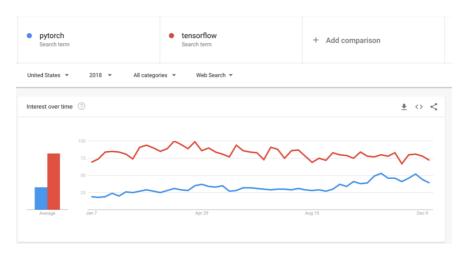
A lot of those advances have been accelerated by the idea of using **language models**, popularized this year by Fast.ai's <u>UMLFit</u> (see also "<u>Understanding UMLFit</u>"). We have then seen other (and improved) approaches like Allen's <u>ELMO</u>, <u>Open AI's transformers</u>, or, more

recently Google's <u>BERT</u>, which beat many SOTA results out of the gate. These models have been described as the "Imagenet moment for NLP" since they provide ready-to-use pre-trained and general models that can be also fine-tuned for specific tasks. Besides language models, there have been plenty of other interesting advances like <u>Facebooks</u> <u>multilingual embeddings</u>, just to name another one. It is interesting to note that we have also seen how quickly these and other approaches have been integrated into more general NLP frameworks such as <u>AllenNLP's</u> or Zalando's <u>FLAIR</u>.



BERT is deeply bidirectional, OpenAl GPT is unidirectional, and ELMo is shallowly bidirectional

Speaking of **frameworks**, this year the "war of the AI frameworks" has heated up. Surprisingly, **Pytorch** seems to be catching up to **TensorFlow** just as <u>Pytorch 1.0 was announced</u>. While the situation around using Pytorch in production is still sub-optimal, it seems like Pytorch is catching up on that front faster than Tensor Flow is catching up on usability, documentation, and education. Interestingly it is likely that the choice of Pytorch as the framework on which to implement the Fast.ai library has played a big role. That being said, Google is aware of all of this and is pushing in the right direction with the <u>inclusion of Keras</u> as a first-class citizen in the framework or the addition of key developer-focused leaders like <u>Paige Bailey</u>. At the end, we all benefit from having access to all these great resources, so keep them coming!



pytorch vs. tensorflow search trends

Interestingly, another area that has seen a lot of interesting developments in the framework space is **Reinforcement Learning**. While I don't think RL research advances have been as impressive as in previous years (only the recent <u>Impala</u> work by DeepMind comes to mind), it is surprising to see that in a single year we have seen all major AI players publish an RL Framework. Google published the <u>Dopamine</u>

<u>framework</u> for research while Deepmind (also inside of Google) published the somewhat competing <u>TRFL</u> framework. Facebook could not stay behind and published <u>Horizon</u> while Microsoft published <u>TextWorld</u>, which is more specialized for training text-based agents. Hopefully, all of this open source goodness will help us see a lot of RL advances in 2019.

Just to finish up on the frameworks front, I was happy to see that Google recently published <u>TFRank</u> on top of Tensor Flow. Ranking is an extremely important ML application that is probably getting less love than it deserves lately.

It might seem like Deep learning has ultimately removed the need to be smart about your data, but that is far from true. There are still very interesting advances in the field that revolve about the idea of improving data. For example, while data augmentation has been around for some time and is key for many DL applications, this year Google published auto-augment, a deep reinforcement learning approach to automatically augment training data. An even more extreme idea is to train DL models with synthetic data. This has been tried in practice for some time and is seen as key to the future of AI by many. NVidia presented interesting novel ideas in their Training Deep Learning with Synthetic Data paper. In our "Learning from the Experts", we also showed how to use expert systems to generate synthetic data that can then be used to train DL systems even after combining with real-world data. Finally, also interesting is the approach of reducing the need to have large quantities of hand-labelled data by using "weak supervision". Snorkel is a very interesting project that aims at facilitating this approach by providing a generic framework.

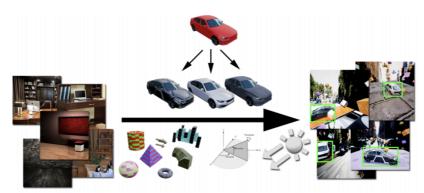


Figure 1. Domain randomization for object detection. Synthetic objects (in this case cars, top-center) are rendered on top of a random background (left) along with random flying distractors (geometric shapes next to the background images) in a scene with random lighting from random viewpoints. Before rendering, random texture is applied to the objects of interest as well as to the flying distractors. The resulting images, along with automatically-generated ground truth (right), are used for training a deep neural network.

As far as more foundational **breakthroughs** in AI, it might be me and my focus, but I haven't seen many. I don't entirely agree with Hinton when he says that this lack of innovation is due to the field having "a few senior guys and a gazillion young guys" although it is true that there is a trend in science where breakthrough research is done at a later age. In my opinion, the main reason for the current lack of breakthroughs is that there are still many interesting practical

applications of existing approaches and variations so it is hard to risk in approaches that might not be practical right away. That is even more relevant when most of the research in the field is sponsored by large companies. In any case, an interesting paper that does challenge some assumptions is "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling". While being highly empirical and using known approaches, it opens the door to uncovering new ones since it proves that the one that is usually regarded as optimal is in fact not. To be clear, I do not agree with Bored Yann LeCun that Convolutional Networks are the ultimate "master algorithm", but rather that RNNs are not either. Even for sequence modeling there is a lot of room for research!. Another highly exploratory paper is the recent NeurIPS best paper award winner "Neural Ordinary Differential Equations", which challenges a few fundamental things in DL including the notion of layers itself.





Spend the holidays with your family, not reading arXiv papers. You'd be wasting your time anyway because—2019 spoiler alert— *CONVOLUTION IS ALL YOU NEED!*

#torched #feelthelearn #Christvolution

7:10 PM - 23 Dec 2018

Interestingly, this last paper was motivated by a project where the authors were looking into **healthcare** data (more concretely Electronic Health Records). I cannot finish this summary without referring to the area of research in the intersection of AI and Healthcare since that is where my focus at Curai is at. Unfortunately, so much is going on in this space that I would need another post only for this. So, I will only point you to the papers that were published at the <u>MLHC conference</u> and the <u>ML4H NeurIPS workshop</u>. Our team at <u>Curai</u> managed to get papers accepted at both, so you will find our papers there among many other interesting ones that should give you an idea of what is going on in our world.