

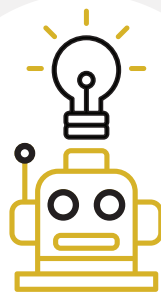
# TOP 4

## DATA SCIENCE TRENDS TO WATCH

### IN 2018

As 2017 comes to a close, it's time to look forward at what's ahead in the wide world of data science. It might not come as a surprise that in 2018, much like 2017, deep learning is the name of the game. But while last year was the year that the idea of deep learning really began to make its way into the mainstream, the coming year will be about how to make deep learning better, faster, and stronger (but not "harder" - in fact, the goal is quite the opposite. Sorry, Daft Punk).

#### Here are our top four data science trends to watch in 2018:



### **AUTOMATED MACHINE LEARNING WILL BECOME A COMMODITY**

Automated machine learning (ML), i.e., the ability to automatically search in the feature transformation and model space, will become a commodity. And that means with this transformation, data science will become less about framework expertise and more about data management and successfully articulating business needs.



3



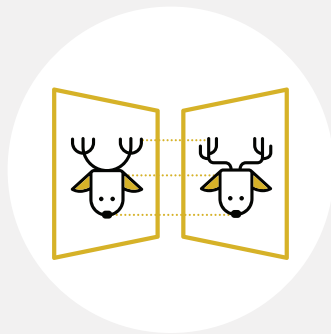
## HARDWARE WILL PLAY A ROLE

For the past several years, the hardware focus in the context of deep learning has largely been on cloud server computing, both for training and execution. But it's becoming increasingly clear that the interest and stakes in deep learning are important enough that it is worth developing specific hardware for it. In 2018, we may very well see more developments in this area.

One specific example of preliminary evidence of this trend is NVIDIA, which has shifted from developing primarily gamers' graphic cards to focusing on optimizing the performance of deep learning and artificial intelligence systems.<sup>1</sup>

This parallels, in many ways, the evolution Bitcoins, the mining of which triggered the design and manufacturing of new hardware architectures that are faster for this specific task.

2



## GENERATIVE ADVERSARIAL NETWORKS (GAN) WILL BECOME MORE MAINSTREAM

GANs are a (relatively) new ML architecture for unsupervised neural networks that have actually been around since 2014, thanks to Ian Goodfellow et al.<sup>2</sup>, and arose out of the need to more easily train models. But they're picking up steam and are starting to become all the rage because, well - they're super useful, and also they have practical applications. But let's back up a bit: what are they exactly, and what makes them so useful?

(1) <https://www.nvidia.com/en-us/data-center/dgx-1/>

(2) <https://arxiv.org/abs/1406.2661>





“[GAN] and the variations that are now being proposed is the most interesting idea in the last 10 years in ML, in my opinion.”

-Yann LeCun, Director of AI Research at Facebook and Professor at NYU <sup>3</sup>



Imagine that you have two people: a professional art imitator and a professional art curator. The job of the former is to create art so impeccably close to an original that the curator is fooled and can't tell the difference. Meanwhile, the latter's role is to be able to successfully weed out pieces that are not made by the original artist based on their expertise and a variety of details.



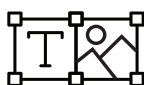
This is, in essence, how GANs work: they are systems of competing neural networks, discriminative and generative. Generative models produce seemingly “natural” data samples, and the discriminative model attempts to distinguish between real and fake data. Every time the generative model fools the discriminative one, it learns, and vice versa.

Normally, neural networks are challenging because they require a lot of data to be trained properly. But these generative models require fewer parameters than normally required for training, so it's more efficient. That's one of the primary reasons this architecture is on the rise.

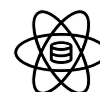
### GAN APPLICATIONS INCLUDE THINGS LIKE:



Image or video generation<sup>4</sup>, including creating extremely high resolution images and image inpainting<sup>5</sup>



Text-to-image synthesis<sup>6</sup>



General ability to train with less data and provide predictions on missing data

Potential pitfalls of GANs include mode collapse<sup>7</sup> and networks that stop learning when they become too strong relative to the other competing network. However, people are beginning to address these challenges, which is setting the stage for showing results in the coming year.

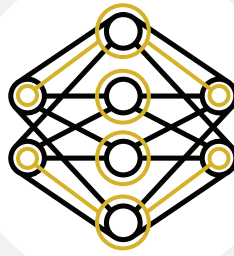
(4) <http://carlvondrick.com/tinyvideo/>

(5) <https://www.cs.toronto.edu/~duvenaud/courses/csc2541/slides/gan-applications.pdf>

(6) <https://arxiv.org/abs/1605.05396>

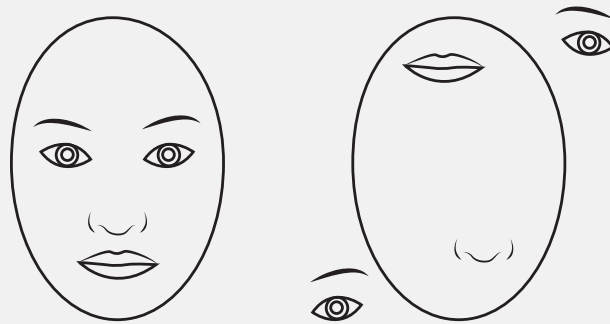
(7) <http://aiden.nibali.org/blog/2017-01-18-mode-collapse-gans/>





## REAL LIFE APPLICATION INNOVATION WILL ACCELERATE

There's no question that innovation in the world of machine learning and deep learning is accelerating - it has been for years. Just one example of a recent development is Geoffrey Hinton's capsule networks<sup>8</sup>, a new kind of neural network architecture that has the potential to spur some interesting new research in the year to come.



*Recent developments, like Hinton's capsule networks, will spur interesting research. But what will be the practical applications? [Source](#).*

**Yet despite all this new research and the acceleration of innovation in the academic space, the fact remains that real people solving real business problems in companies across the world largely struggle to use these cutting-edge innovations in real-life applications.** This is poised to change in 2018 as innovation surrounding real-life applications of machine learning technologies catches up to the theoretical research.

One specific example of a highly-touted innovation that hasn't seen much success in real businesses is sales bots. Bot systems, especially in business-to-customer transactions, are more often than not a hard-coded set of rules. In 2018, these systems will evolve thanks to machine learning and the commoditization of machine learning frameworks trained on real human-to-human conversation.

Wider than just bot applications, global libraries (similar to the transfer learning one widely available today for text and image) will be made available to capture the nature of a conversation and speed up this - and many other - innovations for real-life applications.

(8) <https://arxiv.org/abs/1710.09829>

